

## ALIGNMENTS

RESULT 1  
AAL03248  
ID AAL03248 standard; DNA; 1462 BP

AC AAL03248;

DT 21-NOV-2001 (first entry)

DE Human reproductive system related antigen DNA SEQ ID NO: 5936

KW Human; reproductive system related antigen; reproductive system disorder,

2000

[illegible]XX  
XX  
03-11-2003

XX	17 - TAN-2001 - 2001WS-1501220
DE	

XX 31 - TAN-2000 - 2000115-0170055 DE

PR	04-FEB-2000; 2000US-0180628.
PR	34-FEB-2000; 2000US-0194664

PR	02-MAR-2000; 2000US-0186350.
PR	16-MAR-2000; 2000US-0180974

PR 17-MAR-2000; 2000US-0190076.  
PR 19-APR-2000; 2000US-0198132

PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467

PR 28-JUN-2000; 2000US-0214886;  
PR 30-JUN-2000; 2000US-0215135

PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249219.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249246.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249267.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI: 2001-465570/50.  
XX  
DR  
XX  
PT Isolated nucleic acid molecule encoding a reproductive system antigen  
PT is used in preventing, treating or ameliorating a medical condition -  
XX  
XX  
PS Disclosure; SEQ ID NO 5936; 1297pp + Sequence Listing; English.  
XX  
XX  
CC The present invention provides the protein and coding sequences of a  
CC number of human reproductive system related antigens. These can be used  
CC in the prevention and treatment of reproductive system disorders,  
CC including cancer. The present sequence is a genomic sequence encoding a  
CC protein of the invention.  
XX  
SQ Sequence 1462 BP; 409 A; 271 C; 276 G; 506 T; 0 other;  
Query Match 55.5%; Score 1449.2; DB 22; Length 1462;

Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1454; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

QY 1131 TTGCGTATTTGATTTTCATATGAGCTTTTCTGTAGTTTTCAGAGATTTTAAA 1190
Db 1 TTGCGTATTTGATTTTCATATGAGCTTTTCTGTAGTTTTCAGAGATTTTAAA 60
QY 1191 TCACAGATATATCTAATATATTTAGCCATCAAAAGTTCCAAAAGCCACACCA 1250
Db 61 TCACAGATATATCTAATATATTTAGCCATCAAAAGTTCCAAAAGCCACACCA 120
QY 1251 GTTCCGACTCAATTTGAAGGGTTTTTACAGAGAGGATTAAGTAAAGTGAATTT 1310
Db 121 GTTCCGACTCAATTTGAAGGGTTTTTACAGAGAGGATTAAGTAAAGTGAATTT 180
QY 1311 AATTAAAGCTTAAACCCAGGTAAGAGATAATGAGAAATATTACGGCAATATGAACTGC 1370
Db 181 AATTAAAGCTTAAACCCAGGTAAGAGATAATGAGAAATATTACGGCAATATGAACTGC 240
QY 1371 TTGACTGTTTCTGTGACTTCTCAGCTTAATGTTTAAAGAGGACAAAGAGTTATG 1430
Db 241 TTGACTGTTTCTGTGACTTCTCAGCTTAATGTTTAAAGAGGACAAAGAGTTATG 300
QY 1431 GTGCGATTTAGTAACAGGGTGTGTTTGAATGCTTATTAAGCTCACTTTCCCTGT 1490
Db 301 GTGCGATTTAGTAACAGGGTGTGTTTGAATGCTTATTAAGCTCACTTTCCCTGT 360
QY 1491 TCCTAAGTGTGAATATCTCTTTAACTAGAAAATGCAAAATATTGAACGTATTTTC 1550
Db 361 TCCTAAGTGTGAATATCTCTTTAACTAGAAAATGCAAAATATTGAACGTATTTTC 420
QY 1551 TGTGTAGTATTTAATCTCTTCAATGAGTAATGAAATACCTGTGAGATTAAGAAAT 1610
Db 421 TGTGTAGTATTTAATCTCTTCAATGAGTAATGAAATACCTGTGAGATTAAGAAAT 480
QY 1611 GAGTCTGAGATCTAGTCCCTCTGTGATTCATTAATCTATCCCTTTTCAGTATTA 1670
Db 481 TGTGTAGTATTTAATCTCTTCAATGAGTAATGAAATACCTGTGAGATTAAGAAAT 540
QY 1671 CATGTGCTTAATCTCAGATGAACCATTTTCACCATGCGATTAATCTCTGAGCTT 1730
Db 541 CATGTGCTTAATCTCAGATGAACCATTTTCACCATGCGATTAATCTCTGAGCTT 600
QY 1731 TTCTGGAAATTTGAAGTATCTCTCTTAACCCCAATGTCGAAGGGTATGACTGATCTA 1790
Db 601 TTCTGGAAATTTGAAGTATCTCTCTTAACCCCAATGTCGAAGGGTATGACTGATCTA 660
QY 1791 CCACCTTTGAATTTATGAAACGGGTCAATTTACGAAGTCTGCATGCGTATGAGATATGG 1850
Db 661 CCACCTTTGAATTTATGAAACGGGTCAATTTACGAAGTCTGCATGCGTATGAGATATGG 720
QY 1851 TTTATAGTACAGCTGAGAGATGAACCTCAGCTCCAGATACCATGATGCAACCCAGAT 1910
Db 721 TTTATAGTACAGCTGAGAGATGAACCTCAGCTCCAGATACCATGATGCAACCCAGAT 780
QY 1911 TTTTTCACCTTGATACCGTGCATAGGGAATATTAAGGCGCTGATTTTGTCTATTT 1970
Db 781 TTTTTCACCTTGATACCGTGCATAGGGAATATTAAGGCGCTGATTTTGTCTATTT 840
QY 1971 CCAAGTAAGTATGATTTATCTCTCTTTTATGTTAATGAGAGAAATTTAGCCTCA 2030
Db 841 CCAAGTAAGTATGATTTATCTCTCTTTTATGTTAATGAGAGAAATTTAGCCTCA 900
QY 2031 CTCACAAATGTTCAATTCAGCAAGGCTTCATATCTCTGTGTGGGCTGAGATTAAGAG 2090
Db 901 CTCACAAATGTTCAATTCAGCAAGGCTTCATATCTCTGTGTGGGCTGAGATTAAGAG 960
QY 2091 CTTATTCAGGTTCTGCTCCTAGCTATTAAGCTCAGTTCAGATGCTGAGAGCTGGCTAG 2150
Db 961 CTTATTCAGGTTCTGCTCCTAGCTATTAAGCTCAGTTCAGATGCTGAGAGCTGGCTAG 1020
QY 2151 GGAAGATGATTAATCCTGCTGCTTAACTGAAAAACAGGCTGATCCTGTTAGTATCTA 2210

```

```

Db 1021 GACAGATGATATTCCTGCTGTTACTGAAAAACAGGTGATCTGTACTGATVACTA 1100
QY 2211 TAACTGACCTTAAATGTCACGTTCCTCAATATGACAAATGTTCTAACCAACTAACTCTCA 2270
Db 1081 TAACTGACCTTAAATGTCACGTTCCTCAATATGACAAATGTTCTAACCAACTAACTCTCA 1140
QY 2271 AATGCTTGGAAAGATACTACAAAGCCAAATCTTTATAGAAATTTGGCCCAAGATTAATCTATG 2330
Db 1141 AATGCTTGGAAAGATACTACAAAGCCAAATCTTTATAGAAATTTGGCCCAAGATTAATCTATG 1200
QY 2331 TTGTTTTCATGCTATTTGTTTAAGCTCCAAAGGTTCACTGTGTTTTCGCCCTGTCTGCG 2390
Db 1201 TTGTTTTCATGCTATTTGTTTAAGCTCCAAAGGTTCACTGTGTTTTCGCCCTGTCTGCG 1260
QY 2391 AGTTGTACCCACTGACGTGGGCAAGCTTCTGGGCAATGAGATGAACTGTGTCCTTTT 2450
Db 1261 AGTTGTACCCACTGACGTGGGCAAGCTTCTGGGCAATGAGATGAACTGTGTCCTTTT 1320
QY 2451 CCCACTAACAGTTATCTTTGACTCTGCTGTTATGCTTACAAATGATGATGCTTA 2510
Db 1321 CCCACTAACAGTTATCTTTGACTCTGCTGTTATGCTTACAAATGATGATGCTTA 1380
QY 2511 TGGAAAGCTGTTAAATTAATTTCTGTGTTAAAGAAATTAAGTTTGTATTTTGACA 2570
Db 1381 TGGAAAGCTGTTAAATTAATTTCTGTGTTAAAGAAATTAAGTTTGTATTTTGACA 1440
QY 2571 ATAAACATTAATATATTTTAA 2592
Db 1441 ATAAACATTAATATATTTTAA 1462

RESULT 2
AAL01567
ID AAL01567 standard; cDNA; 779 BP.
AC AAL01567;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen cDNA seq ID NO: 1568.
XX
KW Human: reproductive system related antigen; reproductive system disorder;
XX cancer; gene therapy; ss.
XX
PN Homo sapiens.
XX
PD WO200155320-A2.
XX
02-AUG-2001.
XX
17-JAN-2001; 2001MO-US01339.
XX
31-JAN-2000; 2000US-0179065.
XX
04-FEB-2000; 2000US-0180628.
XX
24-FEB-2000; 2000US-0184664.
XX
02-MAR-2000; 2000US-0186350.
XX
16-MAR-2000; 2000US-0189874.
XX
17-MAR-2000; 2000US-0190076.
XX
18-APR-2000; 2000US-0198123.
XX
19-MAY-2000; 2000US-0205515.
XX
07-JUN-2000; 2000US-0209467.
XX
28-JUN-2000; 2000US-0214886.
XX
30-JUN-2000; 2000US-0215135.
XX
07-JUL-2000; 2000US-0216547.
XX
07-JUL-2000; 2000US-0216880.
XX
11-JUL-2000; 2000US-0217487.
XX
11-JUL-2000; 2000US-0217496.
XX
14-JUL-2000; 2000US-0218290.
XX
26-JUL-2000; 2000US-0220963.
XX
26-JUL-2000; 2000US-0220964.
XX
14-AUG-2000; 2000US-0224518.
XX
14-AUG-2000; 2000US-0224519.
XX
14-AUG-2000; 2000US-0225213.

```

PR	14-AUG-2000	2000US-0225212	2000US-0225265
PR	14-AUG-2000	2000US-0225267	2000US-0225268
PR	14-AUG-2000	2000US-0225269	2000US-0225270
PR	14-AUG-2000	2000US-0225271	2000US-0225272
PR	14-AUG-2000	2000US-0225273	2000US-0225274
PR	14-AUG-2000	2000US-0225275	2000US-0225276
PR	14-AUG-2000	2000US-0225277	2000US-0225278
PR	18-AUG-2000	2000US-0225279	2000US-0225280
PR	22-AUG-2000	2000US-0225281	2000US-0225282
PR	22-AUG-2000	2000US-0225283	2000US-0225284
PR	23-AUG-2000	2000US-0227182	2000US-0227183
PR	23-AUG-2000	2000US-0227189	2000US-0227190
PR	23-AUG-2000	2000US-0227191	2000US-0227192
PR	30-AUG-2000	2000US-0228294	2000US-0228295
PR	01-SEP-2000	2000US-0228297	2000US-0228298
PR	01-SEP-2000	2000US-0228299	2000US-0228300
PR	01-SEP-2000	2000US-0228301	2000US-0228302
PR	01-SEP-2000	2000US-0228303	2000US-0228304
PR	01-SEP-2000	2000US-0228305	2000US-0228306
PR	01-SEP-2000	2000US-0228307	2000US-0228308
PR	05-SEP-2000	2000US-0228309	2000US-0228310
PR	05-SEP-2000	2000US-0228311	2000US-0228312
PR	06-SEP-2000	2000US-0230343	2000US-0230344
PR	06-SEP-2000	2000US-0230345	2000US-0230346
PR	08-SEP-2000	2000US-0231422	2000US-0231423
PR	08-SEP-2000	2000US-0231424	2000US-0231425
PR	08-SEP-2000	2000US-0231426	2000US-0231427
PR	08-SEP-2000	2000US-0231428	2000US-0231429
PR	08-SEP-2000	2000US-0231430	2000US-0231431
PR	08-SEP-2000	2000US-0231432	2000US-0231433
PR	08-SEP-2000	2000US-0231434	2000US-0231435
PR	08-SEP-2000	2000US-0231436	2000US-0231437
PR	12-SEP-2000	2000US-0231968	2000US-0231969
PR	14-SEP-2000	2000US-0233397	2000US-0233398
PR	14-SEP-2000	2000US-0233399	2000US-0233400
PR	14-SEP-2000	2000US-0233399	2000US-0233400
PR	14-SEP-2000	2000US-0233401	2000US-0233402
PR	14-SEP-2000	2000US-0233403	2000US-0233404
PR	14-SEP-2000	2000US-0233405	2000US-0233406
PR	14-SEP-2000	2000US-0233407	2000US-0233408
PR	21-SEP-2000	2000US-0233423	2000US-0233424
PR	21-SEP-2000	2000US-0233425	2000US-0233426
PR	21-SEP-2000	2000US-0233427	2000US-0233428
PR	25-SEP-2000	2000US-0234997	2000US-0234998
PR	25-SEP-2000	2000US-0234997	2000US-0234998
PR	25-SEP-2000	2000US-0234998	2000US-0234999
PR	25-SEP-2000	2000US-0235369	2000US-0235370
PR	29-SEP-2000	2000US-0235370	2000US-0235371
PR	29-SEP-2000	2000US-0235372	2000US-0235373
PR	29-SEP-2000	2000US-0235374	2000US-0235375
PR	29-SEP-2000	2000US-0235376	2000US-0235377
PR	29-SEP-2000	2000US-0235378	2000US-0235379
PR	29-SEP-2000	2000US-0235380	2000US-0235381
PR	29-SEP-2000	2000US-0235382	2000US-0235383
PR	29-SEP-2000	2000US-0235384	2000US-0235385
PR	27-SEP-2000	2000US-0235386	2000US-0235387
PR	27-SEP-2000	2000US-0235388	2000US-0235389
PR	27-SEP-2000	2000US-0235390	2000US-0235391
PR	27-SEP-2000	2000US-0235392	2000US-0235393
PR	27-SEP-2000	2000US-0235394	2000US-0235395
PR	27-SEP-2000	2000US-0235396	2000US-0235397
PR	27-SEP-2000	2000US-0235398	2000US-0235399
PR	27-SEP-2000	2000US-0235400	2000US-0235401
PR	27-SEP-2000	2000US-0235402	2000US-0235403
PR	27-SEP-2000	2000US-0235404	2000US-0235405
PR	27-SEP-2000	2000US-0235406	2000US-0235407
PR	27-SEP-2000	2000US-0235408	2000US-0235409
PR	27-SEP-2000	2000US-0235410	2000US-0235411
PR	27-SEP		

[illegible]

Db 127 ATTTTCGATAGATCCCATTTTGTACAGATTGATGGATGGATAGTAACAAA 186  
 QY 1051 GTAGAACCCATTTAGATTGATTTGATTCACAAATTTGGTTCAAGGCTGGTG 1110  
 Db 187 GTAGAACCCATTTAGATTGATTTGATTCACAAATTTGGTTCAAGGCTGGTA 246  
 QY 1111 ATAGACTATATAACCCAGATTGGCTATTTGATTTGATTCATATGCTTTTCTGTAG 1170  
 Db 247 ATAGACTATATAACCCAGATTGGCTATTTGATTTGATTCATATGCTTTTCTGTAG 306  
 QY 1171 TTTTCAGAGATTTTAAATCAGAGATCATCTAAATGATTTTGGCTATCAAAACT 1230  
 Db 307 TTTTCAGAGATTTTAAATCAGAGATCATCTAAATGATTTTGGCTATCAAAACT 366  
 QY 1231 TCCAAAAGCCACACACACAGTTCCTGACTCAAAATTTGAGGTTTTCAGAGAGGT 1290  
 Db 367 TCCAAAAGCCACACACACAGTTCCTGACTCAAAATTTGAGGTTTTCAGAGAGGT 426  
 QY 1291 AGGATTAAGTAGTGATTTATTTAAAGCTTAAACCTTAGTAGAGATAAGAAATAT 1350  
 Db 427 AGGATTAAGTAGTGATTTATTTAAAGCTTAAACCTTAGTAGAGATAAGAAATAT 486  
 QY 1351 TACGCAATATATGAGACGCTTCACTGTTCTGTGACTTCCATCTAATGTTTAA 1410  
 Db 487 TACGCAATATATGAGACGCTTCACTGTTCTGTGACTTCCATCTAATGTTTAA 546  
 QY 1411 AGAGCAACAACAAAGCTTATGCTGCTTCACTAGTAACACAGGTTGTTTATGATCCCTT 1470  
 Db 547 AGAGCAACAACAAAGCTTATGCTGCTTCACTAGTAACACAGGTTGTTTATGATCCCTT 606  
 QY 1471 ATAGCTCAGTTCCTGCTTCTAAGGTT-GAATAGTCTTAA--CTAGAAAAT 1527  
 Db 607 ATAGCTCAGTTCCTGCTTCTAAGGTTGAAATAGTCTTAAACCTAGGAAAAT 666  
 QY 1528 GCAAAATATTC--AAGTATATTTCTGTGTAGTT--TATTAAGTCTTCAATGAGTGA- 1583  
 Db 667 GCAAAATATTCGAACTGGAATTTTGTGTAGTTAGTCTTCAATGAGTGA- 726  
 QY 1584 --TGATGATTAAGTCTGAGATAGGAAATGAGTTCTGATCTAGTCC 1630  
 Db 727 TGATGATTAAGTCTGAGATAGGAAATGAGTTCTGATCTAGTCC 775  
 RESULT 3  
 AAS94694  
 ID AAS94694 standard; cDNA; 867 BP.  
 AC AAS94694;  
 XX  
 DT 12-MAR-2002 (first entry)  
 DE Rat secreted factor DNA clone P0185\_D11.  
 XX  
 Rat; secreted factor polypeptide; cardiac disease; renal disease; kidney;  
 KM inflammatory disease; congestive heart failure; myocarditis; asthma; ss;  
 KM dilated congestive cardiomyopathy; angina pectoris; cardiac arrhythmia;  
 KM myocardial infarction; pulmonary hypertension; arteriosclerosis; stroke;  
 KM atherosclerosis; cardiac tumor; glomerulonephritis; nephrotic syndrome;  
 KM renal infarction; hereditary nephritis; polycystic kidney disease;  
 KM chronic renal failure; renal vein thrombosis; medullary kidney;  
 KM rheumatoid arthritis; osteoarthritis; psoriasis; restenosis; PDR primer;  
 KM graft versus host reaction; Crohn's disease; ulcerative colitis; probe;  
 KM Alzheimer's disease; gene therapy.  
 XX  
 OS Rattus norvegicus.  
 PN WO200174901-A2.  
 XX  
 PD 11-Oct-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09555.  
 XX  
 PR 31-MAR-2000; 2000US-193548P.

PR 14-MAR-2001; 2001US-0809545.  
 PA (SCIO-) SCIOS INC.  
 XX  
 PI Stanton LW, White RT;  
 XX  
 DR WPI: 2002-010779/01.  
 DR P-PSDB: MAU70147.  
 PT Novel secreted factor polypeptide useful for treating cardiac diseases  
 PT such as arteriosclerosis, myocardial infarction, inflammatory diseases  
 PT such as asthma, stroke, and rheumatoid arthritis and renal diseases  
 XX  
 PS Claim 1; Fig 2; 189pp; English.  
 CC The invention relates to rat secreted factor polypeptides and the  
 CC polynucleotides encoding them. The sequences are useful for treating  
 CC cardiac, renal or inflammatory diseases. These include cardiac diseases  
 CC such as congestive heart failure, myocarditis, dilated congestive  
 CC cardiomyopathy, angina pectoris, myocardial infarction, cardiac  
 CC arrhythmia, pulmonary hypertension, arteriosclerosis, atherosclerosis and  
 CC cardiac tumours, renal diseases such as glomerulonephritis, nephrotic  
 CC syndrome, renal infarction, hereditary nephritis, polycystic kidney  
 CC disease, chronic renal failure, renal vein thrombosis and medullary  
 CC sponge kidney and inflammatory diseases such as asthma, rheumatoid  
 CC arthritis, osteoarthritis, stroke, psoriasis, restenosis, graft versus  
 CC host reaction, Crohn's disease, ulcerative colitis and Alzheimer's  
 CC disease. Sequences AAS94693-AAS94745 represent cDNA clones, which encode  
 CC the secreted factor polypeptides of the invention, and oligonucleotide  
 CC probes and PCR primers.  
 XX  
 SO Sequence 867 BP; 204 A; 202 C; 245 G; 199 T; 17 other:  
 Query Match 19.6%; Score 511; DB 24; Length 867;  
 Best Local Similarity 85.0%; Pred. No. 8.6e-114;  
 Matches 571; Conservative 0; Mismatches 101; Indels 0; Gaps 0;  
 QY 161 GGAAGGCCGTTGCCGCTCGCAAGAAAGCCGTTGACCATGTTGCAACTGGCAG 220  
 Db 172 GAAAGATCCGTTCTCCGCGAGGAAACAGAGCCGTTGACCAATGTTGCAACGGCAG 231  
 QY 221 TTTGAGCAGCAAGAACCCGCGAGCATTTGCAATTTGCTGGAGTGGCTATCACCAGA 280  
 Db 232 TTTGAGCAGTAGAAGACAGGCGCAGCATTTGCAAGTTGCTGGAGTGGCTATCACCAG 291  
 QY 281 GAGCCTGTAAGTATTTTGAATGATGATTTGATTTGCTTCCGTAACCAACCTCAGCA 340  
 Db 292 GAGCTGCTAAGTATTTGATGATGATTTGATTTGCTTCCGTAACCAACCTCAGCA 351  
 QY 341 GGTAAATTTGAGCAATCACTGCGCAGAAATTTGCTTAAATGCTGGAGACTGTCTGC 400  
 Db 352 GGTGCTGTTGAAGAGACAAATGCAATGCAATTTGCTTAAATGCTGGAGACTGTCTGC 411  
 QY 401 CAATCAAGCAAACTAAATCTGTTGCTCAAAAGGCTTGCCTCGAGAACTGACGCA 460  
 Db 412 CAAGTCAAGCAAACTAAATCTGTTGCTCAAAAGGCTTGCCTCGAGAACTGACGCA 471  
 QY 461 GAGAAATTCCTCAAGATGTCCTGCGGCTTCTCAACGAGACCGCTCGGCTTCGAGGTTG 520  
 Db 472 GAGAAATTCCTCAAGATGTCCTGCGGCTTCTCAACGAGACCGCTCGGCTTCGAGGTTG 531  
 QY 521 TGTATGACGAGCACTGGAATTTGAAATGATGTAATAAGCTGATAGATGATGTTG 580  
 Db 532 TGTATGACGAGCACTGGAATTTGAAATGATGTAATAAGCTGATAGATGATGTTG 591  
 QY 581 TGATTTAGCGCTGCTACTTTTGAAGCTTACCTGTTTAAAGCAGAGAACTGCTC 640  
 Db 592 TGATTTAGCGCTGCTACTTTTGAAGCTTACCTGTTTAAAGCAGAGAACTGCTC 651  
 QY 641 ATGACTAGCTTCAGGACTTTTCTTAAAGTAGAGGTCGCTTCCTGCTTTCAGAGAG 700  
 Db 652 CTGACACAGCTCAAGGACTTCTTAAAGTAGAGGTCGCTTCCTGCTTTCAGAGAG 711





PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 25-SEP-2000; 2000US-0234999.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 13-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.

PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 06-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WPI: 2001-465570/50.

DR P-PSDB: AAM94290.

XX Isolated nucleic acid molecule encoding a reproductive system antigen -  
PT is used in preventing, treating or ameliorating a medical condition -

XX Claim 1: SEQ ID NO 261; 1297pp + Sequence Listing; English.

PS The present invention provides the protein and coding sequences of a  
XX number of human reproductive system related antigens. These can be used  
CC in the prevention and treatment of reproductive system disorders,  
CC including cancer. The present sequence is a coding sequence of the  
CC invention.

XX Sequence 415 BP; 114 A; 79 C; 87 G; 131 T; 4 other;

XX Query Match 14.8%; Score 386.4; DB 22; Length 415;

XX Best Local Similarity 97.1%; Pred. No. 9.4e-84; Mismatches 10; Indels 1; Gaps 1;

XX Matches 401; Conservative 1; Mismatches 10; Indels 1; Gaps 1;

QY 2002 TTTATGTTAATGAGAAATTTAGCCTCCACACAAATGTTCAAGCAAGCTTTCA 2061  
DB 1 TTTATGTTAATGAGAAATTTAGCCTCCACACAAATGTTCAAGCAAGCTTTCA 60  
QY 2062 TATCCTGTGCTGGGCTGGGATAGAGAGCTTATTCAGGTTTCGCGCTTACTTACG 2121  
DB 61 TATCCTGTGCTGGGCTGGGATAGAGAGCTTATTCAGGTTTCGCGCTTACTTACG 120  
QY 2122 TCCACTTCACATGCTGGAGAGCTGGGAGGAGACATGTTATTCATTCGCTGGTACTGAA 2181  
DB 121 TCCACTTCACATGCTGGAGAGCTGGGAGGAGACATGTTATTCATTCGCTGGTACTGAA 180  
QY 2182 AAACAGGTGTGATCCTGTTAGTACTATAAGTACCTAAATGCTCACTGTTCAATTTA 2241  
DB 181 AAACAGGTGTGATCCTGTTAGTACTATAAGTACCTAAATGCTCACTGTTCAATTTA 240  
QY 2242 GCAAGTGTCTAACAACTAACTCTTCAATGCTTGGAAAGATCTACCAAGCAACTCT 2301  
DB 241 GCAAGTGTCTAACAACTAACTCTTCAATGCTTGGAAAGATCTACCAAGCAACTCT 300  
QY 2302 TTTATGAATGGGCCAAGTAATCTATGTTGTTGATGGCTATTTAGTACGTCGAAA 2361  
DB 301 TTTATGAATGGGCCAAGTAATCTATGTTGTTGATGGCTATTTAGTACGTCGAAA 360  
QY 2362 GGTTCACGTGTTTTCGCGGCTGCTCGAGTGTTCACACACAGTACGCGCAAG 2414  
DB 361 GGTTCACGTGTTTTCGCGGCTGCTCGAGTGTTCACACACAGTACGCGCAAG 412



RESULT 6  
 AAS24131/c  
 ID AAS24131 standard; cDNA; 446 BP.  
 AC AAS24131;  
 XX  
 XX 07-NOV-2001 (first entry)  
 DE Human ovarian PCR-subtracted cDNA library clone #312.  
 XX  
 XX Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;  
 KM gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;  
 KM primer; probe.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200157207-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 05-FEB-2001; 2001WO-US03733.  
 XX  
 PR 04-FEB-2000; 2000US-0180403.  
 PR 28-MAR-2000; 2000US-0192745.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Algate PA, Mannion J;  
 XX  
 DR WPI; 2001-488879/53.  
 XX  
 PT New polynucleotides encoding ovarian tumour proteins, useful for  
 PT treating ovarian cancer, and as probes, primers, and markers of cancer  
 PT progression -  
 XX  
 PS Example 1; page 161; 378pp; English.  
 XX  
 CC The invention comprises compositions used for the therapy and diagnosis  
 CC of ovarian cancer. The compositions comprise one or more ovarian tumour  
 CC proteins, their associated polynucleotides, or immunogenic portions of  
 CC the proteins. The ovarian tumour polynucleotides and polypeptides are  
 CC useful for stimulating and/or expanding T cells specific for a tumour  
 CC protein. They are also useful for inhibiting the development of cancer in  
 CC a patient with an ovarian tumour DNA or protein by incubating isolated  
 CC T-cells allowing them to proliferate, and administering to the patient.  
 CC The sequences can be used as markers for cancer, for example, to monitor  
 CC ovarian cancer progression. Probes and primers are useful in nucleic acid  
 CC hybridisation, in detecting the presence of complementary sequences in a  
 CC given sample, for preparing mutant species and for preparing other  
 CC genetic constructions. Sequences AAS23820-AAS25231 and AAS25328-AAS25549  
 CC represent human ovarian tumour protein cDNA clones.  
 XX  
 XX Sequence 446 BP; 148 A; 89 C; 82 G; 123 T; 4 other;  
 SQ  
 Query Match 14.0%; Score 364.6; DB 22; Length 446;  
 Best Local Similarity 96.8%; Pred. No. 1.8e-78;  
 Matches 392; Conservative 0; Mismatches 11; Indels 2; Gaps 2;  
 XX  
 QY 2187 GGTGTGATCCTGTTAGTACTATATAGTACCT-AAATGTCACTGTTCAAAATTAGCAA 2245  
 DB 404 GGTGATCCCTCGTGNACTAGTATAGTACCTAAATGTCTGCTCAAAATTAGC-A 346  
 QY 2246 GTGTTTCAACAACTAACTCTTCAATGCTTGGAAAGATCTCAAAAGCAATCTTAT 2305  
 DB 345 GTGTTTCAACAACTAACTCTTCAATGCTTGGAAAGATCTCAAAAGCAATCTTAT 286  
 QY 2306 AGAATGGGCGCAAGATAAATCTATGTTGTTGATGGCTATTGTAAAGCTTCAAGGTT 2365  
 DB 285 AGAATGGGCGCAAGATAAATCTATGTTGTTGATGGCTATTGTAAAGCTTCAAGGTT 226  
 QY 2366 CACTGTGTTTCTGCGCTGTCTGAGTGTGACCACTGACTGCGGCAAGGCTTCTTGGGC 2425  
 DB 225 CACTGTGTTTCTGCGCTGTCTGAGTGTGACCACTGACTGCGGCAAGGCTTCTTGGGC 166

QY 2426 ATGAGTGTAGAACGTGTGCTCTTCCACCTAACAGTTATCTTGACTCTCTTGCCTCTT 2485  
 DB 165 ATGATGTAGAACGTGTGCTCTTCCACCTAACAGTTATCTTGACTCTCTTGCCTCTT 106  
 QY 2486 ATGCTTACAAATGTGATGCTTATGGAAGCGTGTAAATATATTTCTGTTAAAGGA 2545  
 DB 105 ATGCTTACAAATGTGATGCTTATGGAAGCGTGTAAATATATTTCTGTTAAAGGA 46  
 QY 2546 AATTAAGTTTGTCTATTGTTGACAAATTAACATTATATTTT 2590  
 DB 45 AATTAAGTTTGTCTATTGTTGACAAATTAACATTATATTTT 1  
 RESULT 7  
 AAH82706/c  
 ID AAH82706 standard; cDNA; 446 BP.  
 AC AAH82706;  
 XX  
 XX 25-SEP-2001 (first entry)  
 DE Human ovarian tumour associated polynucleotide sequence SEQ ID NO:330.  
 XX  
 XX Human; ovarian tumour; ovarian cancer; diagnosis; gene therapy;  
 KM immunogenic; vaccine; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200151513-A2.  
 XX  
 PD 19-JUL-2001.  
 XX  
 PF 16-JAN-2001; 2001WO-US01575.  
 XX  
 PR 14-JAN-2000; 2000US-0176722.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Algate PA;  
 XX  
 DR WPI; 2001-425866/45.  
 XX  
 PT Novel ovarian tumor proteins, and nucleic acids encoding them, used to  
 PT treat and diagnose cancers, particularly ovarian cancer -  
 XX  
 PS Claim 5; Page 123; 338pp; English.  
 XX  
 CC AAH82377 to AAH83878 represent human ovarian tumour-associated  
 CC polynucleotide sequences which encode ovarian tumour proteins. The  
 CC ovarian tumour protein and polynucleotide sequences have cytostatic  
 CC activity, and can be used in gene therapy and vaccine production. The  
 CC ovarian tumour proteins and polynucleotides can be used to inhibit  
 CC the development of cancer, particularly ovarian cancer. They can also  
 CC be used to diagnose the onset and progression of cancer.  
 XX  
 XX Sequence 446 BP; 148 A; 89 C; 82 G; 123 T; 4 other;  
 SQ  
 Query Match 14.0%; Score 364.6; DB 22; Length 446;  
 Best Local Similarity 96.8%; Pred. No. 1.8e-78;  
 Matches 392; Conservative 0; Mismatches 11; Indels 2; Gaps 2;  
 XX  
 QY 2187 GGTGTGATCCTGTTAGTACTATATAGTACCT-AAATGTCACTGTTCAAAATTAGCAA 2245  
 DB 404 GGTGATCCCTCGTGNACTAGTATAGTACCTAAATGTCTGCTCAAAATTAGC-A 346  
 QY 2246 GTGTTTCAACAACTAACTCTTCAATGCTTGGAAAGATCTCAAAAGCAATCTTAT 2305  
 DB 345 GTGTTTCAACAACTAACTCTTCAATGCTTGGAAAGATCTCAAAAGCAATCTTAT 286  
 QY 2306 AGAATGGGCGCAAGATAAATCTATGTTGTTGATGGCTATTGTAAAGCTTCAAGGTT 2365  
 DB 285 AGAATGGGCGCAAGATAAATCTATGTTGTTGATGGCTATTGTAAAGCTTCAAGGTT 226

Yy	2366	CACNCTGTTTTCGGCCGCTGCTCGTGGAGTTGTCACCACTACACCTGAGGGAAGGCTTCTGGGC	2425		
Db	225	CACGTGTGTTCTTCCGCGCTCTCCTGGAGTTGTACCACTACTGCGGAAGGCTTCTGGGC	166		
Qy	2426	ATGAGTATGAGAACCTTTGTCTCTTTTCCCACTAACAGTATTCCTTGACTCTTGGCTGTT	2485		
Db	165	ATGATATGATGAGAACCTGTGGCTCTTTTCCCACTAACAGTATTCCTTGACTCTCTGGCTGTT	106		
Qy	2486	ATGCTTACAAAAATGGTGTATGGCTTATGGAAGGCTGTTAAATTAATTCCTGTTAAAGGA	2545		
Db	105	ATGCTTACAAAAATGGTGTATGGCTTATGGAAGGCTGTTAAATTAATTCCTGTTAAAGGA	46		
Qy	2546	AATTTAAGTTTGTCTATTTTGTGACAAATAAACATTATATATTTT	2590		
Db	45	AATTTAAGTTTGTCTATTTTGTGACAAATAAACATTATATATTTT	1		
RESULT 8					
XX	AA524329	standard; CDNA: 470 BP.			
XX	AA524329:				
XX	07-NOV-2001	(first entry)			
XX	Human ovarian PCR-subtracted CDNA library clone #510.				
XX	Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;				
XX	gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;				
XX	primer; probe.				
XX	Homo sapiens.				
XX	WO200157207-A2.				
XX	09-AUG-2001.				
XX	05-FEB-2001; 2001WO-US03733.				
XX	04-FEB-2000; 2000US-0180403.				
XX	28-MAR-2000; 2000US-0192745.				
XX	(CORI-) CORIXA CORP.				
XX	Algate PA, Mannion J;				
XX	WPI: 2001-488879/53.				
XX	New polynucleotides encoding ovarian tumour proteins, useful for				
XX	treating ovarian cancer, and as probes, primers, and markers of cancer				
XX	progression				
XX	Example 1; page 194-195; 378pp; English.				
XX	The invention comprises compositions used for the therapy and diagnosis				
XX	of ovarian cancer. The compositions comprise one or more ovarian tumour				
XX	proteins, their associated polynucleotides, or immunogenic portions of				
XX	the proteins. The ovarian tumour polynucleotides and polypeptides are				
XX	useful for stimulating and/or expanding T cells specific for a tumour				
XX	protein. They are also useful for inhibiting the development of cancer in				
XX	a patient with an ovarian tumour DNA or protein by incubating isolated				
XX	T-cells allowing them to proliferate, and administering to the patient.				
XX	The sequences can be used as markers for cancer, for example, to monitor				
XX	ovarian cancer progression. Probes and primers are useful in nucleic acid				
XX	hybridisation, in detecting the presence of complementary sequences in a				
XX	given sample, for preparing mutant species and for preparing other				
XX	genetic constructions. Sequences AA523820-AA52531 and AA525328-AA52549				
XX	represent human ovarian tumour protein CDNA clones.				
XX	Sequence 470 BP; 119 A; 98 C; 96 G; 151 T; 6 other;				
XX	Query Match 11.6%; Score 303.2; DB 22; Length 470;				

Query Match	11.6%	Score 303.2	DB 22	Length 470
Best Local Similarity	97.9%	Pred. No. 1,4e-63		
Matches 327	Conservative 0	Mismatches 5	Indels 2	Gaps 2
OY 1952	CCTGATTTTTTGTCTTATTTCCAACTAGTAGTCATTAATCTCTTCCCTTTTATGTTAA	2011		
Db 2	CCTGATTTTTTGTCTTATTTCCAACTAGTAGTCATTAATCTCTTCCCTTTTATGTTAA	61		
OY 2012	TGAGAGAAATTTAGCCTCCACATCAACAAATGTCCAATTCGCAAGGCTTTCATATCCTTGGCT	2071		
Db 62	TGAGAGAAATTTAGCCTCCACATCAACAAATGTCCAATTCGCAAGGCTTTCATATCCTTGGCT	121		
OY 2072	GTGGATCCGTGATAAGAGAGCTTATTCAGGTTTCCGCCCCTAGTATTAAGTCCATCCATCCAC	2131		
Db 122	GTGGATCCGTGATAAGAGAGCTTATTCAGGTTTCCGCCCCTAGTATTAAGTCCATCCATCCAC	181		
OY 2132	ATGCTGGAGACTGGCGGTAGGAGACAGATGTATTATCATCTGTGTTACTGAAAAACAGGTGT	2191		
Db 182	ATGCTGGAGACCGGCGGTAGGAGACAGATGTATTATCATCTGTGTTACTGAAAAACAGGTGT	241		
OY 2192	GATCCTGTTAGTAGTACTATTAAGTAGCACTTAAATGTCACTGTTCAAATTACCAAGTCTTC	2251		
Db 242	GATCCTGTTAGTAGTACTATTAAGTAGCACTTAAATGT-CACGTGTCAAAATTAGCCAGNG-TG	299		
OY 2252	TACCAACTAAACCTCTTCAAAATGCTTGGAAAGAT	2285		
Db 300	TACCAACTAAACCTCTTCAAAATGCTTGGAAAGAT	333		
RESULT 9				
ID AAH82918	AAH82918 standard; cDNA: 470 BP.			
XX AAH82918:				
XX 25-SEP-2001 (first entry)				
DE Human ovarian tumour associated polynucleotide sequence SEQ ID NO:542.				
XX Human; ovarian tumour; ovarian cancer; diagnosis; gene therapy;				
KM Immunogenic; vaccine; ss.				
XX Homo sapiens.				
PN W0200151513-A2.				
XX 19-JUL-2001.				
XX 16-JAN-2001; 2001WO-US01575.				
XX 14-JAN-2000; 2000US-0176722.				
XX (CORI-) CORIXA CORP.				
XX Algate PA;				
XX WPI; 2001-425866/45.				
XX Novel ovarian tumor proteins, and nucleic acids encoding them, used to				
XX treat and diagnose cancers, particularly ovarian cancer -				
XX Claim 5; Page 164; 338pp; English.				
XX AAH82377 to AAH83878 represent human ovarian tumour-associated				
XX polynucleotide sequences which encode ovarian tumour proteins. The				
XX ovarian tumour protein and polynucleotide sequences have cytosolic				
XX activity, and can be used in gene therapy and vaccine production. The				
XX ovarian tumour proteins and polynucleotides can be used to inhibit				
XX the development of cancer, particularly ovarian cancer. They can also				
XX be used to diagnose the onset and progression of cancer.				
XX Sequence 470 BP; 119 A; 98 C; 96 G; 151 T; 6 other;				

Query Match	11.68; Score 303.2; DB 22; Length 470;
-------------	--

Best Local Similarity 97.98; Pred. No. 1,4e-63;  
Matches 327; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

QY 1952 CCTGATTTTGTGCTTATTCACACTAGTAGATCATATCTTCTGTTTATGTTAA 2011  
|||||  
Db 2 CCTGATTTTGTGCTTATTCACACTAGTAGATCATATCTTCTGTTTATGTTAA 61  
QY 2012 TGAAGAAATTTAGCCTCCACTCAACAATGTTCATTGACAGAGCCTTTCATATCTTGGCT 2071  
|||||  
Db 62 TGAAGAAATTTAGCCTCCACTCAACAATGTTCATTGACAGAGCCTTTCATATCTTGGCT 121  
QY 2072 GTGGGTGGATGATGAGGATTTATTCAGTTCTGCTGCTTACGCTATAGCTCCACTTAC 2131  
|||||  
Db 122 GTGGGTGGATGATGAGGATTTATTCAGTTCTGCTGCTTACGCTATAGCTCCACTTAC 181  
QY 2132 ATGCTGAGAGCTGGCTAGGACAGATGATTCCTGCTGCTTACTGAAGACAGGTGT 2191  
|||||  
Db 182 ATGCTGAGAGCTGGCTAGGACAGATGATTCCTGCTGCTTACTGAAGACAGGTGT 241  
QY 2192 GATCCTGTTAGTACTATATAGTACCTAAATGTCACCTGTTCAATTTAGCAAGTGTTC 2251  
|||||  
Db 242 GATCCTGTTAGTACTATATAGTACCTAAATG-CACCTGTTCAATTTAGCCAGNG-TC 299  
QY 2252 TAACAACTAAACTCTCAATGCTTGAAGAT 2285  
|||||  
Db 300 TAACAACTAAACTCTCAATGCTTGAAGAT 333

RESULT 10  
AAL35611  
ID AAL35611 standard; cDNA; 244 BP.  
XX AAL35611;

DT 08-JAN-2002 (first entry)

DE Human musculoskeletal system related polynucleotide SEQ ID NO 953.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KM anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;  
KM vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KM cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KM neurological disease; infection; human; secreted protein;  
KM musculoskeletal system; ss.

OS Homo sapiens.

PN WO200155367-A1.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01338.

XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216860.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 06-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0234984.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.

PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250310.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 03-JAN-2001; 2001US-0259678.  
XX  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI: 2001-451937/48.  
DR P-PSDB; ABB04029.  
XX  
XX  
PT Isolated polypeptide for treating, preventing and/or prognosing  
PT disorders related to the musculoskeletal system including  
PT musculoskeletal cancers and also for testing and detection e.g.  
PT diagnosis -  
XX  
XX  
PS Claim 1; SEQ ID NO 953; 781bp + Sequence Listing; English.  
XX  
XX  
CC The invention relates to novel genes (AAL34669-AAL37666) and proteins  
CC (AAB03087-ABB04109) associated with the musculoskeletal system useful  
CC for preventing, treating or ameliorating medical conditions e.g. by  
CC protein or gene therapy. The genes are isolated from a range of human  
CC tissues disclosed in the specification. The nucleic acids, proteins,  
CC antibodies and (ant)agonists are useful in the diagnosis, treatment  
CC and prevention of: (a) cancer e.g. breast and ovarian cancer and  
CC other cancers of the adrenal gland, bone, bone marrow, breast,  
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune  
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
CC anemia, autoimmune rheumatoid arthritis, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound  
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;  
CC and (f) infectious diseases such as viral, bacterial, fungal and

CC parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 244 BP; 74 A; 37 C; 44 G; 85 T; 4 other;  
Query Match 9.1%; Score 238.6; DB 22; Length 244;  
Best Local Similarity 98.8%; Pred. No. 4,6e-48;  
Matches 238; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 2370 GTGTTCTGCGCGCTGTCTGAGTTGTACACACTGACGCGCAAGCCTTTGGGCATCG 2429  
DB 1 GTGTTCTGCGCGCTGTCTGAGTTGTACACACTGACGCGCAAGCCTTTGGGCATCG 60  
QY 2430 ATGTGAACTGTGCTTTCCTTCCCACTACATTAATTCCTGCTCCTGCTTATTC 2489  
DB 61 ATGTGAACTGTGCTTTCCTTCCCACTACATTAATTCCTGCTCCTGCTTATTC 120  
QY 2490 TTACAATAATGATGAGCTTATGGAAGGCTGTTAATTAATTCCTGTTAAGCAATTT 2549  
DB 121 TTACAATAATGATGAGCTTATGGAAGGCTGTTAATTAATTCCTGTTAAGCAATTT 180  
QY 2550 AAGTTTGTCTATTTTGTGACAAATTAACATTAATTTTAAAAAATTTTAAAAA 2609  
DB 181 AAGTTTGTCTATTTTGTGACAAATTAACATTAATTTTAAAAAATTTTAAAAA 240  
QY 2610 A 2610  
DB 241 A 241  
RESULT 11  
AAB87098  
ID AAB87098 standard; DNA; 212 BP.  
XX  
XX AAB87098;  
XX  
XX 08-JAN-2001 (first entry)  
XX  
XX DE Rat hepatocyte carcinogenesis biomarker nucleic acid SEQ ID NO:19.  
XX  
XX KW Rat; phenobarbital; carcinogenesis marker; carcinogenesis; detection;  
XX KW identification; carcinogenic; probe; primer; ds.  
XX  
XX OS Rattus norvegicus.  
XX  
XX PN WO200044902-A2.  
XX  
XX PD 03-AUG-2000.  
XX  
XX PF 28-JAN-2000; 2000MO-US00503.  
XX  
XX PR 29-JAN-1999; 99US-0118078.  
XX  
XX PA (SEAR) SEARLE & CO G D.  
XX  
XX PI Bunch RT, Curtis SW, Rodi CP, Morris DL;  
XX  
XX WPI: 2000-505977/45.  
XX  
XX  
PT New nucleic acid encoding a carcinogenic biomarker, induced by  
PT phenobarbital treatment of rat hepatocytes, useful for identifying  
PT carcinogenic compounds -  
XX  
XX  
PS Claim 1; Page 55; 240bp; English.  
XX  
XX  
CC AAB87080 to AAB87656 represent nucleic acid sequences (NI) encoding a  
CC carcinogenesis biomarkers. The carcinogenesis biomarkers are induced by  
CC treating rat hepatocytes with phenobarbital. The nucleic acids are  
CC useful for identifying carcinogenic compounds. The nucleic acid molecules  
CC can be used to derive probes and/or primers for detecting or inducing  
CC carcinogenesis, respectively.

XX Sequence 212 BP; 50 A; 49 C; 62 G; 51 T; 0 other;  
SQ  
Query Match 6.3%; Score 164.8; DB 21; Length 212;  
Best Local Similarity 89.0%; Pred. No. 3.2e-30;  
Matches 178; Conservative 0; Mismatches 22; Indels 0; Gaps 0;  
QY 179 CGCGAAGAAACAGAGCCGTTGACCATGTTGCAACTGGCAGTTTGAGCAGCAAGAACCC 238  
DB 13 CGCGAGGAAACAGAGCCGTTGACCATGTTGCAACGGGAGTTTGAGCAGTAAGAACAC 72  
QY 239 GGGCAGATTTCAGAAATGCTGACATGCTATCACCAGAGACCTGCTAACTGATTT 298  
DB 73 GGGCAGCATTTCAGAGTTGCTGACGCTGCTCACCCTGGGAGTCTGCTAACTGATTT 132  
QY 299 TGACTACTGGATTATGTTGTTCCCTGAACCAACCTCAACGAGGTAATATTGAGGATC 358  
DB 133 CGACTACTGGGATTATGCTCTCCCGAGCCCAACCTCAACGAGGTGCTGTTGAGAGAC 192  
QY 359 AACTTGCCAGAAATTTGGTTA 378  
DB 193 AACATGCCAGAAATTTGGTTA 212  
RESULT 12  
AAC32046  
ID AAC32046 standard; cDNA; 133 BP.  
AC AAC32046;  
XX  
DT 06-OCT-2000 (first entry)  
DE Human secreted protein 5' EST, SEQ ID NO: 36121.  
XX  
KM Human: 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KM gene therapy; chromosome mapping; ss.  
XX  
OS Homo sapiens.  
XX  
PN EP1033401-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 21-FEB-2000; 2000EP-0200610.  
XX  
PR 26-FEB-1999; 9905-0122487.  
XX  
PA (GEST ) GENSET.  
XX  
PI Dumas Milne Edwards J, Duclert A, Giordano J;  
XX  
DR WPI; 2000-500381/45.  
XX  
PS  
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
XX  
XX  
PS Claim 1; SEQ ID 36121; 71pp + CD-ROM; English.  
XX  
CC The present sequence is one of a large number of 5' ESTs derived from  
CC cDNAs encoding secreted proteins. No ORF has yet been conclusively  
CC identified within the present sequence. The 5' ESTs were prepared from  
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
CC derived from the 5' ends of mRNAs and even in those cases where longer  
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.  
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be  
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used  
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC They are used to obtain upstream regulatory sequences and to design  
CC expression and secretion vectors.

XX Sequence 133 BP; 40 A; 18 C; 28 G; 41 T; 6 other;  
SQ  
Query Match 4.2%; Score 109.8; DB 21; Length 133;  
Best Local Similarity 95.2%; Pred. No. 5.5e-17;  
Matches 119; Conservative 4; Mismatches 1; Indels 1; Gaps 1;  
QY 1004 ATCCCATTTTGTACAGAAATG-AATGGATGAAATGATAGCAAAAGTAAGGCCAT 1062  
DB 9 ATCCCATTTTGTACAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 68  
QY 1063 TTGAGTTTACATTGATTCACCAATTTGTTTACGATGAGCTTGATAGACTATATA 1122  
DB 69 TTGAGTTTACATTGATTCACCAATTTGTTTACGAGAGGCTTGTAATAGACTATATA 128  
QY 1123 AACCA 1127  
DB 129 AACCA 133  
RESULT 13  
AAZ41261  
ID AAZ41261 standard; cDNA; 1491 BP.  
AC AAZ41261;  
XX  
DT 18-JAN-2000 (first entry)  
DE Human normal ovarian tissue derived cDNA 40.  
XX  
KM Human: ovary; screening; ovarian cancer; treatment; ss.  
KM  
OS Homo sapiens.  
XX  
PN DE19816395-A1.  
XX  
PD 07-OCT-1999.  
XX  
PF 03-APR-1998; 98DE-1016395.  
XX  
PR 03-APR-1998; 98DE-1016395.  
XX  
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.  
XX  
PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Doll E;  
XX  
DR WPI; 1999-552352/47.  
DR P-PSDB; AAY59820, AAY59821, AAY59822.  
XX  
PT Nucleic acid sequences potentially useful in diagnosis or therapy of  
PT ovarian cancer -  
XX  
XX  
PS Claim 3; Page 153; 274ppp; German.  
XX  
CC This invention describes novel nucleic acid sequences that are highly  
CC expressed in normal ovarian tissue. Artificial chromosomes and cosmid  
CC clones containing the sequences can be used as gene transfer vehicles.  
CC The sequences can be used to produce DNA fragments containing  
CC full-length genes. Host cells transformed with the sequences can be used  
CC to produce polypeptides or polypeptide fragments, which can be used to  
CC screen phage displays for polypeptides that bind to them, or as tools for  
CC identifying agents active against ovarian cancer, or to prepare  
CC medicaments for treating ovarian cancer. The cDNA sequences can be used  
CC to obtain genomic genes, their promoters, enhancers, silencers, exon  
CC structures, intron structures and their splice variants. AAZ41222-241324  
CC represent cDNA sequences derived from normal human ovarian tissue and  
CC which encode the protein fragments represented in AAY59724-Y59837.  
XX  
SQ Sequence 1491 BP; 323 A; 362 C; 455 G; 351 T; 0 other;  
Query Match 2.6%; Score 67.2; DB 20; Length 1491;  
Best Local Similarity 51.2%; Pred. No. 3e-06;  
Matches 206; Conservative 0; Mismatches 193; Indels 3; Gaps 2;

```
QY 351 GAGGAATCAACTTGCCAGAAATTTGGTTAAATGCTGAGAACTGTCTGCCAAATCAAG 410
    || ||| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 194 GATGAACACTTGTTGGCAACCTGATGCAAGCTGCAAGAGAGCTGGCCAGGCGGG 253
QY 411 CAAACTAAACTTGGTTGCTCAAAAGTCTTGTCCCTGAGAAACTGACGACGAAATTTGCT 470
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 254 CTGGGCTCTCGACGCCCTGCGCGGCTGTGATGCTTACGCCAGTTGTAAAGCCAGGTGGGC 313
QY 471 CAAGATGTCCTGGCGCTTTCCTCAACGAGGCCCTGCGGCTTCCGAGTTGTATTGAC 530
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 314 AAGAACTACTGCGCTTCCCTGACGAGCGCTGCGCTGCGGGGGGCGCTGCTGAC 373
QY 531 GTGAACCTTGAATTTGAAATGTATGTAATAAAGCTGAGTAGATTTGTGATTCTGAC 590
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 374 GTGCGCGTGGAGCAAGGGCAAGAGCTGCACAGCGTGGGCGACGTGCATGACCCAGC 433
QY 591 GTCTGACTTACTTTTGACCTTACACTTGTGTTAAGCAGAG-GAACTGCTCATGAGACTAG 649
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 434 CTGGTGCCACCTTCCACGTGACCTCGTGTGCGCTGACTCAGACTGTGGGCCAAG 493
QY 650 CTTCAGGAGACTTTT--CTTTAGTAGAGTGCCTTCTCCTGTTTCAGAGAACTGTCG 707
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 494 ATCCAGGGGCTGTTAGCTCGGCCCACTCTCCCTTCCCTCGCTTACGCAATGCCCTG 553
QY 708 ATCTCAGCTCAGATTTCGACTTGTAAAGAAAAAATTAC 749
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 554 ACGCTGACACTGGCTTCCGAGTCAATCAAGAAAGCTGTAC 595

RESULT 14
AA29824
ID AA29824 standard; cDNA; 1734 BP.
AC AA29824;
XX
XX 12-JUL-2000 (first entry)
DE
DE cDNA sequence of human Hs.111244 polypeptide.
XX
XX Human: pancreatitis-associated protein; PAP: regenerating gene; REG gene;
KW Chronic mucosal injury; Hs.111244 polynucleotide; ulcerative colitis;
KW Crohn's disease; ss.
XX
XX Homo sapiens.
OS
XX WO200014283-A2.
PN
XX 16-MAR-2000.
PD
XX
XX 03-SEP-1999; 99WO-US20098.
PE
XX
XX 04-SEP-1998; 98US-0146969.
PR
XX (UNIM ) UNIV WASHINGTON.
PA
XX Dieckgraefe BK;
XX
XX Dieckgraefe BK;
XX
XX WPI; 2000-257019/22.
DR
XX
XX Diagnosing chronic mucosal injury such as ulcerative colitis and
PT Crohn's disease comprises detecting expression levels of regenerating
PT gene family and a gene represented by a Hs.111244 polynucleotide in a
PT human body sample -
XX
XX
XX Disclosure; Page 40-42; 42pp; English.
XX
XX The present sequence represents the cDNA sequence of human Hs.111244.
CC The Hs.111244 gene is part of the regenerating (REG) gene family.
CC In normal circumstances, REG genes are regionally expressed in low
CC amounts in the small bowel and pancreatic epithelium. However, they
CC are strongly expressed in regions of the colon involved with chronic
CC mucosal injury. The specification describes a method for diagnosing
```

```
CC chronic mucosal injury by detecting expression levels of the REG gene
CC family and a gene represented by a Hs.111244 polynucleotide in a human
CC body sample. The methods are useful for diagnosing chronic mucosal injury
CC such as ulcerative colitis and Crohn's disease by detecting expression
CC levels of the REG gene family and a gene represented by a Hs.111244
CC polynucleotide, respectively, in a human body sample.
XX
XX Sequence 1734 BP; 347 A; 461 C; 536 G; 390 T; 0 other;
SQ
Query Match 2.6%; Score 67.2; DB 21; Length 1734;
Best Local Similarity 51.2%; Pred. No.3.2e-06;
Matches 206; Conservative 0; Mismatches 193; Indels 3; Gaps 2;
QY 351 GAGGAATCAACTTGCCAGAAATTTGGTTAAATGCTGAGAACTGTCTGCCAAATCAAG 410
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 441 GATGAACACTTGTTGGCAACCTGATGCAAGCTGCTGCAAGAGAGCTGGCCAGGCGGG 500.
QY 411 CAAACTAAACTTGGTTGCTCAAAAGTCTTGTCCCTGAGAAACTGACGACGAAATTTGCT 470
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 501 CTGGGCTCTCGACGCCCTGCGCGGCTGTGATGCTTACGCCAGTTGTAAAGCCAGGTGGGC 560
QY 471 CAAGATGTCCTGGCGCTTTCCTCAACGAGGCCCTGCGGCTTCCGAGTTGTATTGAC 530
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 561 AAGAACTACTGCGCTTCCCTGACGAGCGCTGCGCTGCGGGGGGCGCTGCTGAC 620
QY 591 GTGAACCTTGAATTTGAAATGTATGTAATAAAGCTGAGTAGATTTGTGATTCTGAC 590
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 621 GTGCGCGTGGAGCAAGGGCAAGAGCTGCACAGCGTGGGCGACGTGCATGACCCAGC 680
QY 591 GTCTGACTTACTTTTGACCTTACACTTGTGTTAAGCAGAG-GAACTGCTCATGAGACTAG 649
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 681 CTGGTGCCACCTTCCACGTGACCTCGTGTGCGCTGACTCAGACTGTGGGCCAAG 740
QY 650 CTTCAGGAGACTTTT--CTTTAGTAGAGTGCCTTCTCCTGTTTCAGAGAACTGTCG 707
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 741 ATCCAGGGGCTGTTAGCTCGGCCCACTCTCCCTTCCCTCGCTTACGCAATGCCCTG 800
QY 708 ATCTCAGCTCAGATTTCGACTTGTAAAGAAAAAATTAC 749
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 801 ACGCTGACACTGGCTTCCGAGTCAATCAAGAAAGCTGTAC 842

RESULT 15
AACT7444/C
ID AACT7444 standard; cDNA; 1759 BP.
AC AACT7444;
XX
XX 08-FEB-2001 (first entry)
DE
XX Human ORF2999 polynucleotide sequence SEQ ID NO:5997.
XX
XX Human: open reading frame; ORFX; detection; cytosstatic; hepatotropic;
KW vulnerary; antipapillary; antipapillary; antipapillary; antipapillary;
KW anticonvulsant; osteopathic; antipapillary; antipapillary; antipapillary;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antihypertensive;
KW antiviral; antibacterial; antifungal; antineoplastic; antihypertensive;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antineoplastic disease; coagulation;
KW thrombosis; contraceptive; ss.
XX
XX Homo sapiens.
OS
XX WO200058473-A2.
XX
XX 05-OCT-2000.
XX
```

PF 31-MAR-2000: 2000WO-US08621.  
XX  
PR 31-MAR-1999: 9905-0127607.  
PR 02-APR-1999: 9905-0127636.  
PR 05-APR-1999: 9905-0127728.  
PR 30-MAR-2000: 2000US-0540763.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
PI Shinkels RA, Leach M;  
XX  
DR MPI: 2000-602362/57.  
P-PSDB: AAB43235.  
XX  
PT Novel nucleic acids and peptides derived from open reading frame X,  
PT useful for treating e.g. cancers, proliferative disorders,  
PT neurodegenerative disorders and cardiovascular disease -  
XX  
PS Claim 5: Page 5174-5176; 5507pp: English.

CC AAC74446 to AAC7606 encode the proteins given in AAB40237 to AAB43397,  
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;  
CC antiproliferative; antiparkinsonian; nootropic; neuroprotective;  
CC osteoparic; anticonvulsant; antiarthritic; immunosuppressive;  
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
CC antidiabetic; hypotensive; dermatological; immunosuppressive;  
CC antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;  
CC antihypertensive; and antianemic. The sequences can be used for determining  
CC the presence of or predisposition to, or preventing or treating  
CC pathological conditions associated with an ORFX-associated disorder. The  
CC nucleic acids can be used to express ORFX proteins in gene therapy.  
CC vectors. The proteins and nucleic acids may be used to treat cancers,  
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
CC graft vs host disease, cardiovascular disease, diabetes mellitus,  
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
CC nocturnal haemoglobinuria, antineoplastic disease; to enhance  
CC coagulation; to inhibit thrombosis; and as a contraceptive.

CC  
SQ Sequence 1759 BP: 392 A; 548 C; 463 G; 356 T; 0 other;

Query Match 2.6%; Score 67.2; DB 21: Length 1759;  
Best Local Similarity 51.2%; Pred. No. 3.3e-06;

Matches 206; Conservative 0; Mismatches 193; Indels 3; Gaps 2;

OY 351 GAGGAATCAACTGCGCAATTTGGTTAAATGCTGAGAACTGTCGCCAATCAAG 410  
DB 1297 GATGAACACTTGTGTGCAACCTGATGACCTGCTGAGAGAGCTGGCCAGGCCGG 1238  
OY 411 CAACCTAACTGTTGCTCTAAAGGTCCTTGTCTGAGAACTGACGCAAGAAATGCT 470  
DB 1237 CTGGGCTTCGACGCCCTGCGGCCCTGATGCTAGCAAGTTGTAAGCCAGGTGGC 1178  
OY 471 CAAGATGTCCTGGGCTTCTCAACGAGACCTGGGGCTTGGAGTTGTTATGCAC 530  
DB 1177 AAGAACTACTGCGCTGCTACAGGAGCCGTCGGGCTGCGGGGCGCTGCTGAC 1118  
OY 531 GTGAACCTTGAATTAATGATGTAAATAACCTGATAGATTGTTGTTATTTAGC 590  
DB 1117 GTCTGGTGGAGGAGGCAAGAGACTGACAGGCTGGCCAGCTGACACCCAGC 1058  
OY 591 GTGTACTTACTTTTGAAGCTTACCTTGTGTTAAAGCAGA-GAAGCTGATGAGTAG 649  
DB 1057 CTGGTGGCCCACTTCAGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 998  
OY 650 CTTGAGGAGACTTTT--CTTTAGTAGAGTGGCTTCTCTGTTTCAAGAGAACTG 707  
DB 997 ATCCAGGGGCTGTTAGCTCGGCAACTCTCTCTCTCTGCTTCAAGCAATCCCTG 938  
OY 708 ATCTCAGCTCAGGATTTGACTTGTAAAGAAAACTTTAC 749

DB 937 ACCTGAGCACTGGCTTCGAGTATCATCAAGAAAGCTGTAC 896

Search completed: October 13, 2002, 03:59:24  
Job time : 423 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 12, 2002, 21:59:32 ; Search time 77 Seconds  
(without alignments)

Title:	US-09-881-636-1
Perfect score:	2610
Sequence:	1 agccgcgcgcaggtgtgccg .....aaaaaaaaaaaaaaaa 2610

scoring table: IDENT11\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

```

Post-processing: Minimum Match 0%

```
Database : Issued Patents NA:*
```

1:	/cgn2_6/plodata/2/ina/5A_COMB_seq.*
2:	/cgn2_6/plodata/2/ina/5B_COMB_seq.*
3:	/cgn2_6/plodata/2/ina/6A_COMB_seq.*
4:	/cgn2_6/plodata/2/ina/6B_COMB_seq.*
5:	/cgn2_6/plodata/2/ina/6CTUS_COMB_seq.*
6:	/cgn2_6/plodata/2/ina/backfillseq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
	1	67.2	2.6	1734	4	US-09-146-969-5	Sequence 5, App1
	2	50.8	1.9	1677	2	US-08-684-101-1	Sequence 1, App1
	3	50.8	1.9	1677	4	US-09-205-814-1	Sequence 1, App1
	4	41	1.6	448	4	US-08-662-374D-17	Sequence 17, App1
	5	41	1.6	448	4	US-08-952-014C-17	Sequence 17, App1
	6	41	1.6	2409	4	US-09-253-322C-8	Sequence 8, App1
	7	39.6	1.5	841	4	US-09-004-731-40	Sequence 40, App1
	8	39.6	1.5	841	4	US-09-002-731-42	Sequence 42, App1
	9	39.6	1.5	841	4	US-09-032-215-46	Sequence 46, App1
	10	39.6	1.5	841	4	US-09-032-215-48	Sequence 48, App1
	11	39.6	1.5	841	4	US-08-749-699-40	Sequence 40, App1
	12	39.6	1.5	841	4	US-08-749-699-42	Sequence 42, App1
	13	39.4	1.5	198	1	US-08-330-108-16	Sequence 16, App1
	14	39.4	1.5	198	5	PCT-US92-10087-16	Sequence 16, App1
	15	39.4	1.5	6030	1	US-08-441-139-8	Sequence 8, App1
	16	38.6	1.5	3459	2	US-08-960-060-3	Sequence 3, App1
	17	38.6	1.5	3459	4	US-09-307-185-3	Sequence 3, App1
	18	38.6	1.5	3600	3	US-08-894-731-1	Sequence 1, App1
	19	38.6	1.5	6152	4	US-08-973-462-1	Sequence 1, App1
	20	38.4	1.5	403	4	US-09-387-212-2	Sequence 2, App1
	21	38.4	1.5	591	4	US-09-018-635-40	Sequence 40, App1
	22	38.4	1.5	597	2	US-08-800-264A-9	Sequence 9, App1
	23	38.4	1.5	597	2	US-09-018-628-9	Sequence 9, App1
	24	38.4	1.5	597	3	US-09-273-378-9	Sequence 9, App1
	25	38.4	1.5	597	3	US-09-209-605-9	Sequence 9, App1
	26	38.4	1.5	802	4	US-08-896-164-12	Sequence 12, App1
	27	38.4	1.5	960	2	US-08-800-264A-2	Sequence 2, App1

C 28	38.4	1.5	960	2	US -09-018-628-2	Sequence 2,	App11
C 29	38.4	1.5	960	3	US -09-273-378-2	Sequence 2,	App11
C 30	38.4	1.5	960	3	US -09-209-605-2	Sequence 2,	App11
C 31	38.4	1.5	975	2	US -08-600-264A-4	Sequence 4,	App11
C 32	38.4	1.5	975	2	US -09-018-628-4	Sequence 4,	App11
C 33	38.4	1.5	975	3	US -09-273-378-4	Sequence 4,	App11
C 34	38.4	1.5	975	3	US -09-209-605-4	Sequence 4,	App11
C 35	38.4	1.5	1111	2	US -08-600-264A-3	Sequence 3,	App11
C 36	38.4	1.5	1111	2	US -09-018-628-3	Sequence 3,	App11
C 37	38.4	1.5	1311	3	US -09-273-378-3	Sequence 3,	App11
C 38	38.4	1.5	1311	3	US -09-209-605-3	Sequence 3,	App11
C 39	38.4	1.5	1317	2	US -08-600-264A-1	Sequence 1,	App11
C 40	38.4	1.5	1317	2	US -09-018-628-1	Sequence 1,	App11
C 41	38.4	1.5	1317	3	US -09-273-378-1	Sequence 1,	App11
C 42	38.4	1.5	1639	1	US -08-619-103-1	Sequence 1,	App1
C 43	38.4	1.5	1639	1	US -09-018-635-1	Sequence 1,	App1
C 44	38.4	1.5	1639	3	US -09-273-378-1	Sequence 1,	App1
C 45	38.4	1.5	2666	3	US -09-018-635-1	Sequence 1,	App1

## ALIGNMENTS

```

RESULT 1
US-09-146-969-5
; Sequence 5, Application US/09146969
; Patent No. 6228585
; GENERAL INFORMATION:
; APPLICANT: Dieckgraebe, Brian K.
; TITLE OF INVENTION: Gene Markers for Chronic Musculoskeletal Injury
; FILE REFERENCE: 04255, 75314
; CURRENT APPLICATION NUMBER: US/09/146,969
; CURRENT FILING DATE: 1998-09-04
; NUMBER OF SEQ. ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1734
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-146-969-5

```

Query Match	2.68;	Score 67.2;	DB 4;	Length 1734;
Best Local Similarity	51.28;	Pred. No. 8.9e-09;		
Matches 206;	Conservative 0;	Mismatches 193;	Indels 3;	Gaps 2

Oy	351	GAGGAATCACTTGGCCAGAAATTTGGTAAATGCTGGGAACTCTGTCTCCAAATCAAG	410
Db	441	GATGAACACTTGTGTGTGCCAACCTGTATGCAGCTGTGTGGAGAGACCTGTGGCCAGGCCGG	500
Oy	411	CAAACTAAACTTGTGTGTCTCAAAAGTCTCTTGTCCCTGGAACATCGACGACAGAAATGCT	470
Db	501	CTGGGCTCTCGACGGCCCTGGCGGCGTGTGTATGTGCTAGCCAGTTTGTAAAGCAGATGGGC	560
Oy	471	CAGATGTCTCTGTGGGCTTTCCTCAACGAGAGCCCTGGGCTTGGCAGTGTGTATATCAC	530
Db	561	AAAGAACTACATCGCGCTGTGGCTTACAGCAGCGCTGTGGCCCTCGGGGGCGCTCTGGAC	620
Oy	531	GTGAACCTGGAAATTTGAANAATGTATGTATAAAGCGATAGCATTTGTGTGATTTACG	590
Db	621	GTCCTCGTGGAGACAGGCGAAGAGCTGCCACAGCGTGGGCCACGTGGCATCTGCACCCACG	680
Oy	591	GTCGTACCTACTTTTGTGACTTACACTGTGTATTAAGCAGGA-GAAGTCCTATGAGACTAC	649
Db	681	CTGGTGCCACCTTTCACACTGACCTGTGCTGTGGGCGCTGGACATCAGCATCTGTGCCAAG	740
Oy	650	CTTCAGGAGCTTTT--CTTTAGTAGAGGTGTGCTTCTGTGTTTCAGGAGAACTGTG	707
Db	741	ATCCAGGGGCTGTATTAGCTCCGCCAACAATCTCTTCTCCCTGTGCTTACAGCAATGCCCTG	800
Oy	708	ATCTCAGCTCAGGATTTGCAGCTGTTTGAAGAAAACTTTAC	749
Db	801	ACGCTAGCACTGTGCTTCGATCATCAAGAAAGCTGTAC	842

```

CORRESPONDENCE ADDRESS:
ADDRESS: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/205,814
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/684,101
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UPXD:503/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: na
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1677 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: 313..727
US-09-205-814-1

Query Match 1.9%; Score 50.8; DB 4; Length 1677;
Best Local Similarity 61.2%; Pred. No. 0.0003;
Matches 82; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 2477 TTGCCTGTTATGCTTACAAATGCTGATGCGTTATGGAAGCGTGTAAATTAATATTCCT 2536
Db 1542 TTCAATTTTGGCATTGCAATTAACAAGAAATATTTGAATTAATCAGTAAGATGCA 1601
QY 2537 GTTAAGGAATTTAAAGTTGCTATTTTTCACATATAACATTATATTTTAAANAAA 2596
Db 1602 TTTGTAGAAATTTAAATTTTATTATAGACACAACAAGCAAAATAAATACTTTTATAAA 1661
QY 2597 AAAAAAAAAAAAAA 2610
Db 1662 AAAAAAAAAAAAAA 1675

RESULT 4
US-08-642-274D-17/c
Sequence 17, Application US/08642274D
Patent No. 6200749
GENERAL INFORMATION:
APPLICANT: Shlöh, Josef
TITLE OF INVENTION: MUTATED FORMS OF THE ATAXIA-TELANGIECTASIA GENE AND METHOD TO
FILE REFERENCE: 229000033
CURRENT APPLICATION NUMBER: US/08/642,274D
NUMBER OF SEQ ID NOS: 220
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 17
LENGTH: 448
TYPE: DNA

```

```

; ORGANISM: Artificial Sequence
;
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: genomic
US-08-642-274D-17

```

Query Match	1.68;	Score 41;	DB 4;	Length 448;
Best Local Similarity	57.48;	Pred. No. 0.073;		
Matches 74;	Conservative 0;	Mismatches 55;	Indels 0;	Gaps 0;

[illegible]

RESULT 5  
US-08-952-014C-17/c  
; Sequence 17, Application US/08952014C  
; Patent No. 6265158

```

: GENERAL INFORMATION:
: APPLICANT: Shiloh, Yosef
: TITLE OF INVENTION: ATAXIA-TELANGIECTASIA GENE AND ITS
: TITLE OF INVENTION: GENOMIC ORGANIZATION
: NUMBER OF SEQUENCES: 91
: CORRESPONDENCE ADDRESS:

```

STREET: 30500 No. 6265168thwestern Hwy., Suite 410  
CITY: Farmington Hills  
STATE: Michigan  
COUNTRY: U.S.  
ZIP: 48334  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

Query Match	1.68	Score 41	DB 4	Length 448
Best Local Similarity	57.48	Pred. NO.	0.073	
Matches 74	Conservative 0	Mismatches 55	Indels 0	Gaps 0

Oy	1159	TTTTTCGTAGTATTTCAGAGGATTTTTTAAACAGAAATATACTAAATCATATTAG	1218
Db	313	TATTTTCGATRAGTTTGAAGAATTAAGTTCATGTACAGAGCTATTTTATAGAGCTTTTGT	254
Oy	1219	CCATCAAAAGCTCCAAAGCCACACACAGTCCCTGACTCAATTTGAAGGGTTTTT	1278

Db	253	GAGAAACACAGSTTTTAAAGCCCAATGCCAGTTTAAAAAAAATTAACAGGAATTT	194
Qy	1279	AGACACGAG	1287
Db	193	CTAATCTTG	185

RESULT 6  
US-09-293-322C-8  
; Sequence 8, Application US/09293322C

Patent No. 6232110  
GENERAL INFORMATION:  
APPLICANT: Pallas, David C  
APPLICANT: Du, Xianxing  
TITLE OF INVENTION: Coding Sequence for Protein Phosphatase Methyltransferase  
Patent No. 6232110  
TITLE OF INVENTION: Recombinant DNA Molecules and Methods

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	1.68;	Score 41;	DB 4;	Length 2409;
Best Local Similarity	56.28;	Pred. No. 0.18;		
Matches 77;	Conservative 0;	Mismatches 60;	Indels 0;	Caps 0

Qy	2414	CTCTGGCTCTTATAGCTTACAAATAGTGATGCTATGGAAGGCTGTAATTAATTT	2533
	2257	CTCTGCTCTGTCAGCGTCCAGAGTCTGGAGCCAGCTTGTCTTCTTATTATGAT	2316
Qy	2554	CCCTCTTAAGGAATTTAACTTGTCTATTTTGACAAATAAAACNTATATATTTTAAA	2592
	2317	CTTGTTTAAAGAAATATAATCTCCACCTTAAAAAAAAAAAAAAAAAAAAAAA	2376
Qy	2594	AAAAAAAAAAAAAAAAAAAA 2610	
	2377	AAAAAAAAAAAAAAAAAAAA 2393	
Db			

```

1 TELECOMMUNICATION INFORMATION:
2 TELEPHONE: 810-539-5050
3 TELEFAX: 810-539-5055
4 INFORMATION FOR SEQ ID NO: 17:
5 SEQUENCE CHARACTERISTICS:
6 LENGTH: 448 base pairs
7 TYPE: nucleic acid
8 STRANDEDNESS: single
9 TOPOLOGY: linear
10 MOLECULE TYPE: DNA (genomic)
11 US-08-952-014C-17
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
10
```

```

SEQUENCE CHARACTERISTICS:
LENGTH: 448 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-952-014C-17

GENERAL INFORMATION:
APPLICANT: Wu Hunter, Shirley
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
TITLE OF INVENTION: MOLECULES AND USES THEREOF
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:

```

D6	313	TATTTTCGATAGTTTGAAGGAATAGTTCGATCTACAGAGTCATTTTATAGSCCTTTTTC	254	;	COMPUTER READABLE FORM:
				;	MEDIUM TYPE: Floppy disk
QY	1219	CGATATCAAAATTTCCAAAAGCCACACACAGTTCCTGACTCAAAATTTGAAGGGCTTTT	1278	;	COMPUTER: IBM PC compatible
				;	OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/004,731
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/749,699
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 841 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2..728
FEATURE:
NAME/KEY: Xaa = any amino acid
LOCATION: 62
FEATURE:
NAME/KEY: Xaa = any amino acid
LOCATION: 209
US-09-004-731-40

```

```

Query Match          1.5%; Score 39.6; DB 4; Length 841;
Best Local Similarity 67.5%; Pred. No. 0.25;
Matches    54; Conservative   0; Mismatches    26; Indels     0; Gaps      0;

QY    2531 ATTCCTGTTAAAGGAATTAAGTTTGCTATTATTTGGACAATAAACAATATATATTTT 2590
       | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    755 AGTTGTATTAACCTNCAATCGAATTTGTTTTATTATTTGGAAATAAATAATATTTNTGAAA 814
       ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY    2591 AAAAAAAAAAAAAAAAAAAAA 2610
       ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    815 AAAAAAAAAAAAAAAAAAAAA 834

RESULT 8
US-09-004-731-42/c
: Sequence 42, Application US/09004731
: Patent No. 6177258
: GENERAL INFORMATION:
APPLICANT: Wu Hunter, Shirley
APPLICANT: Stiegler, Gary
TITLE OF INVENTION: Gaines, Patrick J.
TITLE OF INVENTION: PLEA PROTEASE PROTEINS, NUCLEIC ACID
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESSES:
ADDRESS: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/004,731
FILING DATE:
CLASSIFICATION:

```

```

: PRIOR APPLICATION DATA: US/08/749,699
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Connell, Gary J.
: REGISTRATION NUMBER: 32,020
: REFERENCE/DOCKET NUMBER: 2618-25-C3
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (303) 863-9700
: TELEFAX: (303) 863-0223
: INFORMATION FOR SEQ ID NO: 42:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 841 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
:
:
: US-09-004-731-42

```

```

Query Match 1.5%; Score 39.6; DB 4; Length 841;
Best Local Similarity 67.5%; Pred. No. 0.25;
Matches 54; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 2531 ATTCCTGTTAAAGAAATTAAGTTGCTATTTTGGACAAATAAATTAATATTTT 2590
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 87 AGTTGATTAACCTNCCATGGAATGTTTATTTTGGAAATAAATATTAATTTTGA 28
      |||||||
QY 2591 AAAAAAAAAAAAAAAAAA 2610
      |||||||
Db 27 AAAAAAAAAAAAAAAAAA 8

RESULT 9
US-09-032-215-46
; Sequence 46, Application US/09032215
; Patent No. 6204010
; GENERAL INFORMATION:
; APPLICANT: Stiegler, Gary L.
; APPLICANT: Gaines, Patrick J.
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES, AND USES THEREOF
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII DOS TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032,215
; FILING DATE: 27-FEB-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 841 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single

```





```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/10087
FILING DATE: 19921118
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/796,066
FILING DATE: 20-NOVEMBER-1991
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: DCI-028PC
TELEPHONE: 617-227-7400
TELEFAX: 617-227-5941
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 198 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single stranded
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: no
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: human
CELL TYPE: T-cell blast
IMMEDIATE SOURCE:
LIBRARY:
CLONE: BD4-T3
PCT-US92-10087-16

```

Query Match 1.5%; Score 39.4; DB 5; Length 198;

Best Local Similarity 56.6%; Pred. No. 0.13; Mismatches 56; Indels 0; Gaps 0;

```

Matches 73; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 2482 TGTATGCTTAAAGTGTGATGCTATGGAAGCTGTTAAATTAATTAATTCCTGTTAA 2541
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 70 TATTTTGTATTAAGTCCTTCCTTATTTGTAAGCTGTTAATTAATTAATTAATTAATTA 129
QY 2542 AGCAATTAAGTGTGCTATTTTGCACATTAATTAATTAATTAATTAATTAATTAATTA 2601
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 130 ATATATTAAAGCAAAATGTTTCAGAAAAAATTAATTAATTAATTAATTAATTAATTAAT 189
QY 2602 AAAAAAAA 2610
DB 190 AAAAAAAA 198

```

## RESULT 15

```

US-08-441-139-8
Sequence 8, Application US/08441139
Patent No. 5773245
GENERAL INFORMATION:
APPLICANT: Wiltup, Dr. Karl D.
APPLICANT: Robinson, Anne S.
TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,139
FILING DATE: 15-MAY-1995

```

```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/089,997
FILING DATE: 06-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8646
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 6030 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1004..4753
US-08-441-139-8

```

Query Match 1.5%; Score 39.4; DB 1; Length 6030;

Best Local Similarity 51.4%; Pred. No. 0.84; Mismatches 86; Indels 0; Gaps 0;

```

Matches 91; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
QY 2433 TAGAAGCTGTGCTTTCCCACTAAGATTTTGTGACCTGCTGCTGTTATGCTTA 2492
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 4810 TTGATTTGCTTCACCCCTGATTCCTTCCTTTGTTTATTTCTCGGAACATTA 4869
QY 2493 CAAATGCTGATGCTTATGGAAGCTGTTAAATTAATTAATTCCTGTTAAAGAAATTA 2552
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 4870 ACAAGCTGATGATTTTGTACTTCTCTTTAATCCAGCTCCGGTAAGCATCATCA 4929
QY 2553 GTTGTCTATTTTACACATAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2609
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 4930 ATTTTATTTGACCTGTTAAGTCATTAATTTTTCCTCCAAAAATTAATTAATTA 4986

```

Search completed: October 13, 2002, 04:01:03

Job time : 98 secs





2 others

## ORIGIN

Query Match 35.1%; Score 915.8; DB 10; Length 1066;  
 Best Local Similarity 95.6%; Pred. No. 7.5e-151;  
 Matches 985; Conservative 0; Mismatches 38; Indels 7; Gaps 4;

```

OY 344 AATATTGAGAAATCACTGCGCAGAAATTTGTTAAATGCTGAGAACGTCTGTCCAA 403
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 11 AATGTCGACACAGCGTCGCGCAGAAATTTGTTAAATGCTGAGAACGTCTGTCCAA 70
OY 404 ATCAAGCAAACTAACTGTTGTTGCTCAAGGCTTGTCCCTGAGAAAAGTGCAGAG 463
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 71 ATCAAGCAAACTAACTGTTGTTGCTCAAGGCTTGTCCCTGAGAAAAGTGCAGAG 130
OY 464 AATGCTCAAGATGCTCGGCTTCTCAACGAGACCTCGGCTTGCAGAGTGTGT 523
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 131 AATGCTCAAGATGCTCGGCTTCTCAACGAGACCTCGGCTTGCAGAGTGTGT 190
OY 524 TATGCACTGAACTTGGAAATTTGAAATGTATGTAAGAAAGCTGATGATGTGTGA 583
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 191 TATGCACTGAACTTGGAAATTTGAAATGTATGTAAGAAAGCTGATGATGTGTGA 250
OY 584 TTTGACGCTGACCTGACCTTTGAGCTTACCTGTGTTAAGACAGAACCTGCTCATG 643
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 251 TTTGACGCTGACCTGACCTTTGAGCTTACCTGTGTTAAGACAGAACCTGCTCATG 310
OY 644 GACTACCTTCAGGAGACTTTTCTTTAGTAGAGGTCGCTTCTCTGTTTCAGAGAAC 703
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 311 GACTACCTTCAGGAGACTTTTCTTTAGTAGAGGTCGCTTCTCTGTTTCAGAGAAC 370
OY 704 TCTGATCTTCAGCTCAGGATTTTCGACTTTGTAAGAAAACTTACTCAGTATGGAAC 763
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 371 TCTGATCTTCAGCTCAGGATTTTCGACTTTGTAAGAAAACTTACTCAGTATGGAAC 430
OY 764 AACAGTATGGAAGGCTCTTAAAGAAATATATTAAGATTAATGATGATGATGATG 823
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 431 AACAGTATGGAAGGCTCTTAAAGAAATATATTAAGATTAATGATGATGATGATG 490
OY 824 GTAACAACTATTCAGCTACGCTAAGTCAATTTGATGTTGCCCCAGCTGCCCCTAATA 883
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 491 GTAACAACTATTCAGCTACGCTAAGTCAATTTGATGTTGCCCCAGCTGCCCCTAATA 550
OY 884 AGAAGCCCAATGTAGTCTTTTCTTTCTGTGTTTCACATTCATGACATGACGCTA 943
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 551 AGAAGCCCAATGTAGTCTTTTCTTTCTGTGTTTCACATTCATGACATGACGCTA 610
OY 944 ACAGAGCTGATTTTCTGCGCTTGGAGAGGATTCAAATAGTGTAGATTTTCTGCATG 1003
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 611 ACAGAGCTGATTTTCTGCGCTTGGAGAGGATTCAAATAGTGTAGATTTTCTGCATG 670
OY 1004 ATCCCAATTTTGTACAGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 1063
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 671 ATCCCAATTTTGTACAGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 730
OY 1064 TGAGTTTACATTTGATTTTCCACAATTTGTTGAGTGTGAGTGTGATGATGATGAT 1123
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 731 TGAGTTTACATTTGATTTTCCACAATTTGTTGAGTGTGAGTGTGATGATGATGAT 790
OY 1124 ACCAGATTTGCTAATTTGATTTTATGAGCTTTTCTGCTAGATTTTCTGAGAGATT 1183
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 791 ACCAGATTTGCTAATTTGATTTTATGAGCTTTTCTGCTAGATTTTCTGAGAGATT 850
OY 1184 TTTTAAATCAGAGAAATCAATGATGATGATGATGATGATGATGATGATGATGATG 1242
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 851 TTTTAAATCAGAGAAATCAATGATGATGATGATGATGATGATGATGATGATGATG 910
OY 1243 CACACACATTTCTGACCTCAAAATTTGAA--GGGTTTTTACACAGAGAGGATGATG 1297
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 911 CACACACATTTCTGACCTCAAAATTTGAAAGGGGTTTTTACACAGAGAGGATGATG 970
OY 1298 AGTAGTAGTAAATTAAGAGTTAACCTGAGTAGAGTAAT--GAGAAATTTTACGGG 1356
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 971 AGTAGTAGTAAATTAAGAGTTAACCTGAGTAGAGTAATGAGAAATTTTACGGG 1030
  
```

OY 1357 AATAATGAA 1366  
 ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 DB 1031 AATAATGGA 1040

RESULT 2  
 B1560922 938 bp mRNA linear EST 05-SEP-2001  
 LOCUS 603254306F1 NIH\_MGC\_97 Homo sapiens CDNA clone IMAGE:5296656 5',  
 DEFINITION mRNA sequence.  
 ACCESSION B1560922  
 VERSION B1560922.1 GI:15448236  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 1 (bases 1 to 938)  
 NIH-MGC <http://mgi.nhl.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shih-Wei  
 Toshiyuki and Piero Carninci (RIKEN)  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLM11750 row: 1 column: 01  
 High quality sequence stop: 748.

FEATURES  
 source 1..938  
 location/Qualifiers  
 1..938  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5296656"  
 /clone\_lib="NIH\_MGC\_97"  
 /lab\_host="DH10B"  
 /note="organ: testis; Vector: pBluescriptR (modified  
 pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI (glucag  
 ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTTTTT-3',  
 size-selected for average insert size 2.2 kb and  
 normalized to ROT 5. This is a primary library enriched  
 for full-length clones and constructed using the  
 cap-trapper method (carninci, in preparation). Library  
 constructed by M. Brownstein (NIH/NHGRI, National  
 Institutes of Health). Note: this is a NIH\_MGC Library."

BASE COUNT 253 a 189 c 208 g 287 t 1 others  
 ORIGIN

Query Match 27.5%; Score 716.8; DB 10; Length 938;  
 Best Local Similarity 98.8%; Pred. No. 5.7e-116;  
 Matches 732; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

```

OY 217 GCAGTTTGAGCAGCAAGACCGGCGCAGCATTTCAAGATTTGCTGAGACTGTGCTATCACC 276
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 4 GAGGTTTGAGCAGCAAGACCGGCGCAGCATTTCAAGATTTGCTGAGACTGTGCTATCACC 63
OY 277 CAGAGAGCTGCTAAGTATTTGACTACTGGGATTTATGTTGTTCTGCAACCAACTGA 336
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 64 CAGAGAGCTGCTAAGTATTTGACTACTGGGATTTATGTTGTTCTGCAACCAACTGA 123
OY 337 ACAGAGTAATTTGAGGATCAACTTGGCAGAAATTTGTTAAATGCTGAGAACTGTC 396
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 124 ACAGAGTAATTTGAGGATCAACTTGGCAGAAATTTGTTAAATGCTGAGAACTGTC 183
OY 397 TGTCCAAATCAAGCAAACTGTTGCTCAAGAGTCTGTTGCTGAGAAACTGA 456
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 184 TGTCCAAATCAAGCAAACTGTTGCTCAAGAGTCTGTTGCTGAGAAACTGA 243
  
```



AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: CLONETECH Laboratories, Inc.  
 cDNA Library Preparation: CLONETECH Laboratories, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: L10M1212 row: k column: 04  
 High quality sequence stop: 687.

FEATURES  
 source  
 Location/Qualifiers  
 1..723  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:433787"  
 /clone\_lib="NIH\_MGC\_81"  
 /lab\_host="DH10B (TI phage-resistant)"  
 /note="Organ: muscle (skeletal); Vector: pDNR-LIB  
 (Clontech); Site:1: Sfil (ggccgcctggcc); Site:2: Sfil  
 (ggccatctggcc); 5' and 3' adaptors were used in cloning  
 as follows: 5' adaptor sequence:  
 5'-ATCTAGAGCGCGAGCGCGCACATG-dT(30)BN-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size  
 1.55 kb (range 1.0-4.0 kb). 15/15 colonies contained  
 inserts by PCR. This library was enriched for full-length  
 clones and was constructed by Clontech Laboratories (Palo  
 Alto, CA)."

BASE COUNT 201 a 147 c 137 g 238 t  
 ORIGIN  
 Query Match 25.2%; Score 658.8; DB 10; Length 723;  
 Best Local Similarity 97.8%; Pred. No. 9e-106;  
 Matches 710; Conservative 0; Mismatches 12; Indels 4; Gaps 4;  
 QY 1584 TGATGAATACCTGTGAGATAGAGAAATGAGTCTGAGATCTGCTGATTCAGT 1643  
 DB 1 TGATGAATACCTGTGAGATAGAGAAATGAGTCTGAGATCTGCTGATTCAGT 60  
 QY 1644 TAGTAATCTATCCTCTTTTACGATTAATGATGCTTAATCTAGATGAACCAATTTACCA 1703  
 DB 61 TAGTAATCTATCCTCTTTTACGATTAATGATGCTTAATCTAGATGAACCAATTTACCA 120  
 QY 1704 TGGCAGTGTATCTCATCTCTGGGCTTTCTGGGAATGAAGATCTCTTAAACCCA 1763  
 DB 121 TGGCAGTGTATCTCATCTCTGGGCTTTCTGGGAATGAAGATCTCTTAAACCCA 180  
 QY 1764 ATTGTCAAGGATAGTACTGTATCTACACCTTTGAATTAATGAACGGGCAATTTAGC 1823  
 DB 181 A-TGTCAAGGATAGTACTGTATCTACACCTTTGAATTAATGAACGGGCAATTTAGC 239  
 QY 1824 AAGTCTGATGCGTATGAGATATGCTTTATATGTAACACCTTAAGAGAAATGAACTACCG 1883  
 DB 240 AAGTCTGCA-TGGCTATGAGATATGCTTTATATGTAACACCTTAAGAGAAATGAACTACCG 298  
 QY 1884 TCCAGATACCATGATGACACCAAGATTTTTCACCTGGATGATCTGACCTAGGGAAT 1943  
 DB 299 TCCAGATACCATGATGACACCAAGATTTTTCACCTGGATGATCTGACCTAGGGAAT 358  
 QY 1944 AATAAAGGCTGATTTTGTCTTATTCACCTAAGTAGATATATCTTCTCTTTT 2003  
 DB 359 AATAAAGGCTGATTTTGTCTTATTCACCTAAGTAGATATATCTTCTCTTTT 418  
 QY 2004 TATGTTAATGAGAATTTAGCTCCACCTCAACAATGTTCAATTCAGCAAGGCTTTCA 2063  
 DB 419 TATGTTAATGAGAATTTAGCTCCACCTCAACAATGTTCAATTCAGCAAGGCTTTCA 478  
 QY 2064 TCCTGCTGTGGGTGTGATAGAGGCTTATTCAGGTTTCCCTGCTATTAATGCTC 2123  
 DB 1 TCCTGCTGTGGGTGTGATAGAGGCTTATTCAGGTTTCCCTGCTATTAATGCTC 2123

DB 479 TCCTGCTGTGGGTGTGATAGAGGCTTATTCAGGTTTCCCTGCTATTAATGCTC 538  
 QY 2124 CACTTCACATGCTGAGACTGGCGTAGGACAGATGATATTCATCTGTTACTGAAAA 2183  
 DB 539 CACTTCACATGCTGAGACTGGCGTAGGACAGATGATATTCATCTGTTACTGAAAA 598  
 QY 2184 ACAGGTGTATCCTGTTAGTACTATATAGTGACCTAAATGTCACGTTCAAATTTAGC 2243  
 DB 599 ACAGGTGTATCCTGTTAGTACTATATAGTGACCTAAATGTCACGTTCAAATTTAGC 658  
 QY 2244 AAGTGTTCACAACTAACTTCAATGCTTGAAGATACCTCAAGCAATCTTT 2303  
 DB 659 CAGTGTTCACAACTAACTTCAATGCTTGAAGATACCTCAAGCAATCTTT 716  
 QY 2304 ATAGAA 2309  
 DB 717 TTGAAA 722  
 RESULT 5  
 LOCUS BG122146 733 bp mRNA linear EST 30-JAN-2001  
 DEFINITION 602349588F1 NIH\_MGC\_90 Homo sapiens cDNA clone IMAGE:444653 5',  
 mRNA sequence.  
 ACCESSION BG122146  
 VERSION BG122146.1 GI:12615655  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 733)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: L10M10220 row: o column: 22  
 High quality sequence stop: 675.  
 Location/Qualifiers  
 1..733  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:444653"  
 /clone\_lib="NIH\_MGC\_90"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: liver; Vector: pCMV-SPORT6; Site:1: NotI;  
 Site:2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 1.7 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

BASE COUNT 203 a 135 c 164 g 231 t  
 ORIGIN  
 Query Match 24.9%; Score 650; DB 10; Length 733;  
 Best Local Similarity 97.1%; Pred. No. 3.1e-104;  
 Matches 715; Conservative 0; Mismatches 15; Indels 6; Gaps 5;  
 QY 292 GTGATTTGACTACTGAGATATGTTGCTGCAACCA-CCTCAACAGGTAATTT 350  
 DB 1 GTGATTTGACTACTGAGATATGTTGCTGCAACCA-CCTCAACAGGTAATTT 60  
 QY 351 GAGGATCAACTGCGCAGAAATTTGTTAAATCTGGAGAACTGTCTGCAAAATCAAG 410  
 DB 61 GAGGATCAACTGCGCAGAAATTTGTTAAATCTGGAGAACTGTCTGCAAAATCAAG 120

QY 411 CAAACTAACTGTGTCCTCAAGGTCCTGTCCTGAGAACTGACGAGAACTTCT 470  
 DB 121 CAACTAACTGTGTCCTCAAGGTCCTGTCCTGAGAACTGACGAGAACTTCT 180  
 QY 471 CAAGATGTCCTGCGGCTTCTCTCAAGGAGCCGCGGCTGGAGGTTGTATGAC 530  
 DB 181 CAAGATGTCCTGCGGCTTCTCTCAAGGAGCCGCGGCTGGAGGTTGTATGAC 240  
 QY 531 GTGAACCTTGAATAATGATGTAATAAAGCTGATAGATTTGTGATTTAGC 590  
 DB 241 GTGAACCTTGAATAATGATGTAATAAAGCTGATAGATTTGTGATTTAGC 300  
 QY 591 GTGCTACTACTTTTGTAGCTTACACTGTGTTTAAAGCAAGAACTGCTGATC 650  
 DB 301 GTGCTACTACTTTTGTAGCTTACACTGTGTTTAAAGCAAGAACTGCTGATC 360  
 QY 651 TTGAGGACTTTTCTTTAGAGAGTCGCTTCTCTGTTTCAAGAGAACTGATC 710  
 DB 361 TTGAGGACTTTTCTTTAGAGAGTCGCTTCTCTGTTTCAAGAGAACTGATC 420  
 QY 711 CTCAGCTCAGGATTTGCACTTTGTTAAGAAAACTTTACTCAGATTGGAACAAGCTG 770  
 DB 421 CTCAGCTCAGGATTTGCACTTTGTTAAGAAAACTTTACTCAGATTGGAACAAGCTG 480  
 QY 771 ATTGAAGGTCCTTAAAGGAAAAATATATTAAGATTATTTCAATGTTGGTATTAAC 830  
 DB 481 ATTGAAGGTCCTTAAAGGAAAAATATATTAAGATTATTTCAATGTTGGTATTAAC 540  
 QY 831 TATTCAGCTAGTCAGTAAAGTCATTTGTAAGTTGCCCCACCTGCTTAATTAAGAAC 890  
 DB 541 TATTCAGCTAGTCAGTAAAGTCATTTGTAAGTTGCCCCACCTGCTTAATTAAGAAC 599  
 QY 891 CCAATATGATCTTTTCTTTCTTCTGTTTACATTTACAGCACTGCACTAAAGGCT 950  
 DB 600 CCAATATGATCTTTTCTTTCTTCTGTTTACATTTACAGCACTGCACTAAAGGCT 658  
 QY 951 GATTTTCTGCGCTTGGAGAAGTATCAAAATAGTATGATTTTCTGCAATGATCCAT 1010  
 DB 659 GATTTTCTGCGCTTGGAGAAGTATCAAAATAGTATGATTTTCTGCAATGATCCAT 715  
 QY 1011 TTTTGTACAGAAATGA 1026  
 DB 716 TTTTGTACAGAAATGA 731

RESULT 6  
 BE86599 763 bp mRNA linear EST 20-OCT-2000  
 LOCUS 601509345F1 NIH\_MGC\_71 Homo sapiens cdna clone IMAGE:3910762 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BE86599  
 VERSION BE86599.1 GI:10341048  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 763)  
 AUTHORS NIH-MGC http://mgc.ncl.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.lnl.gov  
 Plate: LLM8726 row: 3 column: 11  
 High quality sequence stop: 701.

FEATURES  
 source Location/Qualifiers  
 1..763  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3910762"  
 /clone\_lib="NIH-MGC\_71"  
 /tissue\_type="leiomysarcoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: uterus; Vector: pCMV-Sport6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 2.1 kb.  
 BASE COUNT 181 a 183 c 201 g 197 t 1 others  
 ORIGIN  
 Query Match 24.8%; Score 647; DB 10; Length 763;  
 Best Local Similarity 96.1%; Pred. No. 1e-103;  
 Matches 697; Conservative 0; Mismatches 20; Indels 8; Gaps 3:  
 QY 41 GCTGCTCTGGGCTCTGACGTACCGAGTCCCGAGAGAGGAGTCTTAAGAGCA 100  
 DB 1 GCTGCTCTGGGCTCTGACGTACCGAGTCCCGAGAGAGGAGTCTTAAGAGCA 60  
 QY 101 CCTGCTGTGTTTACCCGAGAGCGGAGCGGAGCCCGGAACTCTGGCAGCGCTCA 160  
 DB 61 CCTGCTGTGTTTACCCGAGAGCGGAGCGGAGCCCGGAACTCTGGCAGCGCTCA 120  
 QY 161 GGAAGGCGCTTCCGCTCGGAGAAACAGAGCCGTTGACCATGTTGCACTGGCAG 220  
 DB 121 GGAAGGCGCTTCCGCTCGGAGAAACAGAGCCGTTGACCATGTTGCACTGGCAG 180  
 QY 221 TTGAGCAGCAAGAACCGGCGCAGCATTTCAAGTTGCTGACCTGGCTATNCCCGA 280  
 DB 181 TTGAGCAGCAAGAACCGGCGCAGCATTTCAAGTTGCTGACCTGGCTATNCCCGA 240  
 QY 281 GAGCCTGTAAGTATTTTGACTACTGAGATTATGTTTCTCTGAACCAACTCAAGA 340  
 DB 241 GAGCCTGTAAGTATTTTGACTACTGAGATTATGTTTCTCTGAACCAACTCAAGA 300  
 QY 341 GGTATATTTAGAGAAATCACTTGGCAGAAATTTGGTTAAATGCTGAGAACTGCTG 400  
 DB 301 GGTATATTTAGAGAAATCACTTGGCAGAAATTTGGTTAAATGCTGAGAACTGCTG 360  
 QY 401 CAAATCAAGCAAACTAACTTGGTTCGCAAGAGTCCCTGCTGAGAAACTGAGCA 460  
 DB 361 CAAATCAAGCAAACTAACTTGGTTCGCTGCAAGAGTCCCTGCTGAGAAACTGAGCA 420  
 QY 461 GAGAATTCCTCAAGATGTCGCGGCTTCTCTCAACGAGAGCCGCTTGCAGGTTG 520  
 DB 421 GAGAATTCCTCAAGATGTCGCGGCTTCTCTCAACGAGAGCCGCTTGCAGGTTG 480  
 QY 521 TGTATATGACGTGAACCTTGAATAATGTATGTAAAAAGCTGGATTAGATTGCTG 580  
 DB 481 TGTATATGACGTGAACCTTGAATAATGTATGTAAAAAGCTGGATTAGATTGCTG 540  
 QY 581 TGATTCTAGCTGCTGACTACTTTTGAAGCTTACACTTGTGTTAAGCAGAGAACTGCTC 640  
 DB 541 TGATTCTAGCTGCTGACTACTTTTGAAGCTTACACTTGTGTTAAGCAGAGAACTGCTC 600  
 QY 601 ATGACACTGCTTCAAGGAACTTTTCTTTAGTAGAGTCCGCTGCTGCTGCTCAGG 660  
 DB 661 ATGACACTGCTTCAAGGAACTTTTCTTTAGTAGAGTCCGCTGCTGCTGCTCAGG 620  
 QY 699 AGAAGCTGATCCGAGCTGAGATTTGCACTTGTTAAGAAAAAATTACTGATGTT 758  
 DB 715 AGAAGCTGATCCGAGCTGAGATTTGCACTTGTTAAGAAAAAATTACTGATGTT 714  
 QY 759 GGAAC 763  
 DB 715 GGACC 719

RESULT 7  
 BE613468

LOCUS	BE613468	556 bp	mRNA	linear	EST 20-OCT-2000
DEFINITION	601504234P1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3905984 5', mRNA sequence.				
ACCESSION	BE613468				
VERSION	BE613468.1	GI:9895065			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strassberg, Ph.D. Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a> Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LHAM9714 row: c column: 09 High quality sequence stop: 643. Location/Qualifiers 1..656 /organism="Homo sapiens" /db_xref="taxon:9606" /clone IMAGE:3905984" /clone_lib="NIH_MGC_71" /tissue_type="leiomysarcoma" /lab_host="DH10B (phage-resistant)" /note="Organ: uterus; Vector: pCMV-SPORT6; Site:1; NotI; Site:2; SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.1 kb." Average insert size 2.1 kb.				
BASE COUNT	160 a 153 c 184 g 159 t				
ORIGIN					
Query Match	23.8%; Score 622.2; DB 10; Length 656;				
Best Local Similarity	98.6%; Pred. NO. 2.4e-99;				
Matches 638; Conservative	0; Mismatches 8; Indels 1; Gaps 1;				
QY	22 GAGGGGTAGCAGGGGTGCGGTGCTGCTGGGGGTCTGCAGGTCCAGGAAGA 81				
DB	1 GAGGGGTGAGCAGGGGTGCGGTGCTGCTGGGGGTCTGCAGGTCCAGGAAGA 60				
QY	82 GGGGACTCTTAGAAGCCACTGCTGTGTATTACCGGAGGAGCGGCGGCCCGC 141				
DB	61 GGGGACTCTTAGAAGCCACTGCTGTGTATTACCGGAGGAGGAGCGGCGGCCCGC 120				
QY	142 GAACCTCTGGCAGCGCTCAGAAAGGCGCTGCGCTCGCGAAGAAACAGCCGTTGA 201				
DB	121 GAACCTCTGGCAGCGCTCAGAAAGGCGCTGCGCTCGCGAAGAAACAGCCGTTGA 180				
QY	202 CCATGGTTTCACTGGCACTTTTGAGCAGCAAAACCCGGCCAGCATTTTCAGATTGCTGG 261				
DB	181 CCATGGTTTCACTGGCACTTTTGAGCAGCAAAACCCGGCCAGCATTTTCAGATTGCTGG 240				
QY	262 ACTGTGGCTATCACCCAGAGAGCGTCTTAAGTGAATTTTGACTACTTGGAATTAATGTTGTTTC 321				
DB	241 ACTGTGGCTATCACCCAGAGAGCGTCTTAAGTGAATTTTGACTACTTGGAATTAATGTTGTTTC 300				
QY	322 CTGAACCCCACTCAACAGAGGTAATATTGTTGAGGAATCAACTTGGCCAGAAATTTGGTTAAAA 381				
DB	301 CTGAACCCCACTCAACAGAGGTAATATTGTTGAGGAATCAACTTGGCCAGAAATTTGGTTAAAA 360				
QY	382 TGCCTGAGAACTGTGTGTCCAATCAAAAGCAAACTTAACCTGTTGCTCAAGAGTCCTGG 441				
DB	361 TGCCTGAGAACTGTGTGTCCAATCAAAAGCAAACTTAACCTGTTGCTCAAGAGTCCTGG 420				
QY	442 TCCCTGAGAACTGAAGCAGAGAAATTTGCTCAAGATGTCTCGGGCTTCTCAACGGAGC 501				

Db	421	TGCGTAGAAGAACTGACCCAGAGAATTTGGTCAAGATGTCTGGGGCTTCCTCAACGAGGC	480
Oy	502	CCTGCGGCTTGCGAGGTTGTGTATTATGACAGCTGAACCTTGAAATTTGAATGTATGTAAAA	561
Db	481	CTGCGGGCTTGGAGAGTTGTGTATTATGACAGCTGAACCTTGAAATTTGAATGTATGTAAAA	540
Oy	562	AGCTGATAGAGATTTGTGTGATTTCTAGCGTGTACTACCTTTTGAGCTTTACATTGCT	621
Db	541	AGCTGATAGAGATTTGTGTGATTTCTAGCGTGTACTACCTTTTGAGCTTTACATTGCT	599
Oy	622	TTAACGAGAGAACTGCTCATAGACTTAACCTTCAGGACTTTTTCTTT	668
Db	600	TTAACGAGAGAACTGCTCATAGACTTAACCTTCAGGACTTTTTCTTT	646
RESULT 8			
LOCUS	BG428930	700 bp	mRNA
DEFINITION	602500527F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4614127 5'		linear EST 14-MAR-2001
ACCESSION	BG428930		mRNA sequence.
VERSION	BG428930.1	GI:13335436	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/		
JOURNAL	1 (bases 1 to 700)		
COMMENT	Unpublished (1999) National Institutes of Health, Mammalian Gene Collection (MGC) Contact: Robert Strausberg, Ph.D. Email: rga@nih.gov Tissue Procurement: CLONTECH Laboratories, Inc. cDNA Library Preparation: CLONTECH Laboratories, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMLN at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> Plate: LCM1365 row: e column: 08 High quality sequence stop: 690.		
FEATURES			
Source	Location/Qualifiers		
	1..700		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4614127"		
	/clone_lib="NIH_MGC_75"		
	/lab_host="DH10B (TI phage-resistant)"		
	/note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site_1:		
	SfiI (ggcgccctgcgccc) Site_2: SfiI (ggccatagccc); 5' and		
	3' adaptors were used in cloning as follows: 5' adaptor		
	sequence: 5'-CACGAGCCATTATGGCC-3' and 3' adaptor sequence:		
	5'-TTCTAGAGCGCAGCGGCCGACATG-drr(30)BN-3' (where B = A,		
	C, or G and N = A, C, G, or T). Average insert size 1.65		
	kb (range 0.5-4.0 kb). 15/15 colonies contained inserts		
	by PCR. This library was enriched for full-length clones		
	and was constructed by Clontech Laboratories (Palo Alto,		
	CA). Note: this is a NIH_MGC Library."		
BASE COUNT	204 a 130 c 136 g 230 t		
ORIGIN			
	Query Match 23.8%; Score 621.8; DB 10; Length 700;		
	Best Local Similarity 97.4%; Pred. NO. 2.8e-99;		
	Matches 685; Conservative 0; Mismatches 12; Indels 6; Gaps 5;		
Oy	1210	GATATTGACCATCAAACTCTCAAAGCCACACACCAGTTCCTGACTCAATTTGA	1269
Db	1	GATATTGACCATCAAACTCTCAAAGCCACACACCAGTTCCTGACTCAATTTGA	60
Oy	1270	AGGTTTTAGACAGAGAGGTAGATTAGTAGTGATTTAATTAAAGCTTAACCCTAG	1329
Db	61	AGGTTTTAGACAGAGAGGTAGATTAGTAGTGATTTAATTAAAGCTTAACCCTAG	120

Oy	1330	GTAAAGAGTAAATGACAAATATTATACGGCAATATGGAACGTGCACCTGTTCTTGCGTAC	1389
Oy	1390	TTCTCTCACTCTAATGTTTAAAGAGCGCAACAAACCTATGATGGCCATTTCAGTAACAC	1449
Oy	1450	GGTGTGTTTATAGATGCCCTTATATAGCTCAGTTTCCCTGTTCTTAAAGTGTGAATACTG	1509
Oy	241	GGTGTTGTTTAAAGACCTTATATAAGCGTCAGTTTCCCGTTCCTTAAGTGTGAATACTG	300
Oy	1510	TCTTTAAACTAGAAAAAAGCAAAATATTTGAACGTGATTTCTGCTGTGTATTTATTACTG	1569
Oy	301	TCTTTAAACTAGAAAAAATGCAAAATATTGAACTGATATTT--TGCTGTATGTTGATTACTG	359
Oy	1570	TTCCATGATGTAATGATGATATACCTGTGAGGATATGGAATAGATTCTCAGATCAGTCC	1629
Oy	360	TTCCATGATGTAATGATGATATACCTGTGAGGATATGGAATATGATTCTCAGATCAGTCC	419
Oy	1630	CTCTGTGATTTCACTTAGTAATCTATCCTCTTTTCAGTATTTACATGTGCTTAACTCAGAT	1689
Oy	420	CTCTGTGATTTCACTTAGTAATCTATCCTCTTTTCAGTATTTACATGTGCTTAACTCAGAT	479
Oy	1690	GAACCATTTCAACATGAGGAGGAGTGTTCATCTGCATCTCGGGCTTTTCGGGAATTTGAAGTATC	1749
Oy	480	GAACCATTTCAACATGAGGAGGAGTGTTCATCTGCATCTCGGGCTTTTCGGGAATTTGAAGTATC	539
Oy	1750	TCTCTTAAACCCCAATTTGTCAAGGGTAGTAGCTGTATCTACCACTTTTGAATTT--ATTGA	1807
Oy	540	TCTCTTAAACCCCAATTTGTCAAGGGTAGTAGCTGTATCTACCACTTTTGAATTTATTGA	599
Oy	1808	AACGGGTAATTTACGAAGTCTGCATTGGCTATGTGAGATATGTTTATATAGTACAGCTTAC	1867
Oy	600	ACCAGGTCAGTTTACGAAGTCTGCATTGG--TATGAGATATGTTTATATAGTACAG--CTAG	657
Oy	1868	AGAATGAACACTCACCGTCCAGA--TAACATGCGATCCACCACCA	1909
Oy	658	AGAATGAACACTCACCGTCCAGTATTAACCATGATCCACCACCA	700
RESULT	9		
LOCUS	BF969068	926 bp	mRNA
DEFINITION	6022270005f1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4358112 5',		
ACCESSION	BF969068		
VERSION	BF969068.1	GI:12336283	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 926)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgarbes-remail.nih.gov		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: Life Technologies, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.lnl.gov		
	plate://Image995 row: j column: 01		
	High quality sequence stop: 672.		
FEATURES	location/qualifiers		
	1..926		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		

Query Match	Best Local Similarity	Score	DB	Length	gaps
Matches 691: Conservative	96.2%	Pred. No. 1.3e-97	0: Mismatches 20: Indels 7: Gaps 6:		
QY 1543	GATATTCTGCTGTAGATTATTACTCTTCCATTGAGTAATGATGAATACCTGTGAGGA	1602			
DB 29	GATATTCTGCTGTAGATTATTACTCTTCCATTGAGTAATGATGAATACCTGTGAGGA	88			
QY 1603	TAGGAATAGTTCGAGATCTAGTCCCTGCTGATCTAGTAATCTATCCCTCTTT	1662			
DB 89	TAGGAATAGTTCGAGATCTAGTCCCTGCTGATCTAGTAATCTATCCCTCTTT	148			
QY 1663	CAGTATTACATGTGCTTAATCTCAGATGAACCATTTCCACATGGCAGTATTCTCATCT	1722			
DB 149	CAGTATTACATGTGCTTAATCTCAGATGAACCATTTCCACATGGCAGTATTCTCATCT	208			
QY 1723	CTGGGCTTTCTCGGGAATTTGAAGTATCTCTCTTAACCCCAATTGTCAAGGTAAGTACT	1782			
DB 209	CTGGGCTTTCTCGGGAATTTGAAGTATCTCTCTTAACCCCAATTGTCAAGGTAAGTACT	268			
QY 1783	GTATACCTACCATTTGGAATTTATGAAGGGGTCATTTACGAAGCTCATTTGGGCTATGG	1842			
DB 269	GTATACCTACCATTTGGAATTTATGAAGGGGTCATTTACGAAGCTCATTTGGGCTATGG	328			
QY 1843	AGATATGTTTATAGTACAGCTTAGAAGATGAACCTCACCGTCCAGATTAACCATCATGC	1902			
DB 329	AGATATGTTTATAGTACAGCTTAGAAGATGAACCTCACCGTCCAGATTAACCATCATGC	388			
QY 1903	ACCCAGATTTTTCACACCTTGATACCTGTCACTAGGGAATTAATAAGCCCTGATTTTTT	1962			
DB 389	ACCCAGATTTTTCACACCTTGATACCTGTCACTAGGGAATTAATAAGCCCTGATTTTTT	448			
QY 1963	GTCTTATTCGAAGTATGATCAATTAATCTCTCTCTTTTATGTATAGAGAAATTT	2022			
DB 449	GTCTTATTCGAAGTATGATCAATTAATCTCTCTCTTTTATGTATAGAGAAATTT	508			
QY 2023	ACCCGCCAATCAACATGTTCATTCAGCAAGCTTTTATATCCCTGTGTCGTGCG	2082			
DB 509	ACCCGCCAATCAACATGTTCATTCAGCAAGCTTTTATATCCCTGTGTCGTGCG	568			
QY 2083	ATAAGAGAGTTATTTAGG-TTTTCGTGCCCTTACT-ATTAGTCCCA-CTTCCACATCTGGA	2139			
DB 569	ATAAGAGAGTTATTTAGG-TTTTCGTGCCCTTACT-ATTAGTCCCA-CTTCCACATCTGGA	628			
QY 2140	GACTGGCGTAGGAGCAGATGTATTCATCTGTGTACTGAAAAACAGGTGATCTGT	2199			
DB 629	GACTGGCGTAGGAGCAGATGTATTCATCTGTGTACTGAAAAACAGGTGATCTGT	686			
QY 2200	TAGTGTACTTAAAGAGACCTAAATGTCACTGTTCCAATTGCAAGGTCTTATCAAA	2257			
DB 687	TAGTGTACTTAAAGAGACCTAAATGTCACTGTTCCAATTGCAAGGTCTTATCAAA	742			

BASE COUNT 232 a 201 c 200 g 293 t

ORIGIN

Query Match 23.4% Score 612; DB 10; Length 926;  
Best Local Similarity 96.2% Pred. No. 1.3e-97;  
Matches 691: Conservative 0: Mismatches 20: Indels 7: Gaps 6:

/clone="IMAGE:4358112"  
/clone\_lib="NIH\_MGC\_84"  
/tissue\_type="adrenal cortex carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/Note="Organ: adrenal gland; Vector: PCMV-SPOB6; Site: 1; Notif. Site: 2; Salt: Cloned unidirectionally; oligo-dt primed. Average insert size 1.229 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC library."

SOURCE	ORGANISM	human.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE	Nih-Mgc http://mgc.nci.nih.gov/.	
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)	
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgrabb@email.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLNL1218 row: g column: 20 High quality sequence start: 4 High quality sequence stop: 607. Location/Qualifiers	
FEATURES	source	
	1..1211	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="IMAGE:5533483"	
	/clone_lib="NIH_MGC_71"	
	/tissue_type="telomysarcoma"	
	/lab_host="DH10B (phage-resistant)"	
	/note="Organ: uterus; Vector: pCMV-SPORT6; Site.1: NotI; Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb."	
BASE COUNT	269 a 355 c 305 g 281 t	1 others
ORIGIN		
Query Match	23.3%; Score 609.2; DB 10; Length 1211;	
Best Local Similarity	98.6%; Pred. No. 3.7e-97;	
Matches	625; Conservative 0; Mismatches 8; Indels 1; Gaps 1	
OY	2 GCCGGCGCAGGCTGGCCGGGAGGAGGGGTGACAGGGTGCCGTGCTGGGGCTGCAG	61
Db	30 GTCCGCCGCGAGGGTGGCCGGGAGGAGGGGTGACAGGGTGCCGTGCTGGGGCTGCAG	89
OY	62 GTACCAGATGCCCAAGAGAGGAGACTCTAAGAAGCCACGTCGTGTTAAACCGGGA	121
Db	90 GTACACGAGATGCCCAAGAGAGGAGACTCTAAGAAGCCACGTCGTGTTAAACCGGGA	149
OY	122 GCGAGCGCGCAGCCCCCGCGAATCTCTGGCAGCGCTCAGAAAGGCCCTGGCCCTGCG	181
Db	150 GCGAGCGCGCAGCCCCCGCGAATCTCTGGCAGCGCTCAGAAAGGCCCTGGCCCTGCG	209
OY	182 GAAGGAACAAGAGCGGTGACCATGTTGCCACTGGCAGTTGAGCAGAAAGAACCCGGC	241
Db	210 GAAGGAACAAGAGCGGTGACCATGTTGCCACTGGCAGTTGAGCAGAAAGAACCCGGC	269
OY	242 CAGATTTCAGAAATGCTGAGACTGTGGCTATCACCCAGAGAGCCTGCTAAGTAGTTTTGA	301
Db	270 CAGATTTCAGAAATGCTGAGACTGTGGCTATCACCCAGAGAGCCTGCTAAGTAGTTTTGA	329
OY	302 CTACATGGGATTAATGTTGTTCTCGTAACCCAACTCAACGAGGATTAATTTGAGAGATCAAC	361
Db	330 CTACATGGGATTAATGTTGTTCTCGTAACCCAACTCAACGAGGATTAATTTGAGAGATCAAC	389
OY	362 TTGCCAGAAATTTGTTAAATACTGTGAGAACTGTCTGTCCTCAATCAACCAAACTTAAC	421
Db	390 TTGCCAGAAATTTGTTAAATACTGTGAGAACTGTCTGTCCTCAATCAACCAAACTTAAC	449
OY	422 TGGTTGCTCAAAGGTCTTTGCTCCCTGAGAACTGACGACGAGAAATTTGCTCAAGATGCT	481
Db	450 TGGTTGCTCAAAGGTCTTTGCTCCCTGAGAACTGACGACGAGAAATTTGCTCAAGATGCT	509
OY	482 GCGGGTTTCTCAAGGAGGCCCTGGGGCTGGAGAGTTGTTATTCACAGTGAACCTTGA	541
Db	510 GCGGGTTTCTCAAGGAGGCCCTGGGGCTGGAGAGTTGTTATTCACAGTGAACCTTGA	569

QY	542	AAATGAAATGATGATGATAAAAGCTGGATGAGATTGCTGATGCTGACGATGCTA-       	600
DB	570	AATTGAAATGATGATGATAAAAGCTGGATGAGATTGCTGATGCTGACGATGCTA-       	629
QY	601	CTTTTGAGCTTACACTGTGTGTTTAAAGCAGAGGAA       	634
DB	630	CTTTTGAGCTTACACTGTGTGTTTAAAGCAGAGGAA       	663
RESULT 11			
LOCUS	BM472975	1435 bp mRNA	EST 05-FEB-2002
DEFINITION	AGENCOURT.6462582 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5519576		
ACCESSION	5', mRNA sequence.		
VERSION	BM472975		
KEYWORDS	BM472975.1 GI:18522017		
SOURCE	EST.		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 1435)		
JOURNAL	NIH-MGC <a href="http://mhc.nci.nih.gov/">http://mhc.nci.nih.gov/</a> .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cga@bbs-rcmail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN). DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> Plate: LNLML182 row: d column: 09 High quality sequence start: 53 High quality sequence stop: 454. Location/Qualifiers 1. 1435 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_image="5519576" /clone_id="NIH_MGC_71" /issue_type="telomysarcoma" /lab_host="DH10B (phage-resistant)" /note="Organ: uterus; Vector: pCMV-Sport6; Site.1: NotI; Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb." 416 a 338 c 264 g 416 t 1 others		
BASE COUNT			
ORIGIN			
Query Match	23.3%	Score 609.2; DB 10; Length 1435;	
Best Local Similarity	93.7%;	Pred. No. 3.5e-97;	
Matches 635; Conservative	0; Mismatches 43; Indels 0; Gaps 0;		
QY	919	TTTCATTCATTCATGCACTGACAGCTGAACAGGCTGATTTCTGCGCTTTGGAGAAGTGAATC       	978
DB	77	TTTACATTCATTCATGCACTGACAGCTGAACAGGCTGATTTCTGCGCTTTGGAGAAGTGAATC       	136
QY	979	AAATATGATGATGATTTTCTGCAATAGATCCCATTTTGTACAGAAATTTGAATGGAATGCAAT       	1038
DB	137	AAATATGATGATGATTTTCTGCAATAGATCCCATTTTGTACAGAAATTTGAATGGAATGCAAT       	196
QY	1039	AGGTAAAGCAAAAGTGAAGCCCATTTGAGTTTACATTTGATTCACAAATTTGGTTTCAG       	1098
DB	197	AGGTAAAGCAAAAGTGAAGCCCATTTGAGTTTACATTTGATTCACAAATTTGGTTTCAG       	256
QY	1099	GTAGAGCTTGATGATGACATATTAACACAGATTTGCTATTTTGAATTTTCATATGCGCTTT       	1158
DB	257	GTAGAGCTTGATGATGACATATTAACACAGATTTGCTATTTTGAATTTTTCATATGCGCTTT       	316
QY	1159	TTTTTCTGTAACTTTTCAGAGAGATTTTAAATCAAGAAATCACTAAATGATATTTAG       	1218



```

Db 317 TTTTCTCTAGTTTCTAGAGATTTTAAATACAGATCACTAATGATTTAG 376
Oy 1219 CCTATCAAAACCTCCAAAGGCCACACCAGTTCCTGACTCAAAATTTGAAGGTTTTT 1278
Db 377 CCTATCAAAACCTCCAAAGGCCACACCAGTTCCTGACTCAAAATTTGAAGGTTTTT 436
Oy 1279 AGACAGAGAGGTAGATTAAGTAGTACGTTTAAAGTTAAAGCTTAACCTAGCTAAGATA 1338
Db 437 AGACAGAGAGGTAGATTAAGTAGTACGTTTAAAGTTAAAGCTTAACCTAGCTAAGATA 496
Oy 1339 AATGAGAAATATTAGGCAATTAATGAAGTCTTCTGACTGTTTCTGGTACTTCTACT 1398
Db 497 AATGAGAAATATTAGGCAATTAATGAAGTCTTCTGACTGTTTCTGGTACTTCTACT 556
Oy 1399 CTAAATGTTTAAAGAGCAACAAAGCTTATGTCGATTCAGTAACGACGAGTGTGT 1458
Db 557 CTAAATGTTTAAAGAGCAACAAAGCTTATGTCGATTCAGTAACGACGAGTGTGT 616
Oy 1459 TTAGATGCTTTTATACCTGATTCCTGTTCTTAAAGTGTGAATACGCTTTTAAAC 1518
Db 617 TTAGATGCTTTTATACCTGATTCCTGTTCTTAAAGTGTGAATACGCTTTTAAAC 676
Oy 1519 TAGAAAAATGCAAAATTTGACATGATTTCTGTCGTAGTATTATTAACCTTCCATTGA 1578
Db 677 TAGAAAAATGCAAAATTTGACATGATTTCTGTCGTAGTATTATTAACCTTCC 736
Oy 1579 GTGAATGATGATATCCTG 1596
Db 737 ATTGAAGCGAATGATG 754

```

```

RESULT 12
LOCUS BE613690 609 bp mRNA linear EST 20-OCT-2000
DEFINITION 60150487F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3906261 5',
ACCESSION BE613690
VERSION BE613690.1 GI:9895287
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 609)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
Plate: LHAM9714 row: n column: 22
High quality sequence stop: 607.
Location/Qualifiers
1. 609
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3906261"
/clone_lib="NIH_MGC_71"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: uterus; Vector: pCMV-SPORT5; Site: J. NCI;
Site 2: Sall; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."

```

BASE COUNT

148 a 144 c 172 g 145 t

Query Match 23.3%; Score 607.4; DB 10; Length 609;  
 Best Local Similarity 99.8%; Pred. No. 9.6e-97;  
 Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Oy 22 GAGGGGTGAGAGAGGTGCGCTGCTGGGCTGAGGCTGACCGAGTCCCGAGAGA 81
Db 1 GAGGGGTGAGAGAGGTGCGCTGCTGGGCTGAGGCTGACCGAGTCCCGAGAGA 60
Oy 82 GGGGACTCTTAAGAACCCACCTGCTGTTTACCCGCGACGAGCCGAGCCCGC 141
Db 61 GGGGACTCTTAAGAACCCACCTGCTGTTTACCCGCGACGAGCCGAGCCCGC 120
Oy 142 GAATCTCTGGAGGGCTGAGGAAGGCGCTGCGCTGCGGAGGAACAGAGCCGTGA 201
Db 121 GAATCTCTGGAGGGCTGAGGAAGGCGCTGCGCTGCGGAGGAACAGAGCCGTGA 180
Oy 202 CCATGTTTGCACATGCGGCTTGGAGCAGCAAGAACCCGCGAGCATTTTACAGATTGCTGG 261
Db 181 CCATGTTTGCACATGCGGCTTGGAGCAGCAAGAACCCGCGAGCATTTTACAGATTGCTGG 240
Oy 262 ACTGTGCTATCACCCAGAGAGCCGTGTAAGTATTTTACTGAGGATTAATGTTTC 321
Db 241 ACTGTGCTATCACCCAGAGAGCCGTGTAAGTATTTTACTGAGGATTAATGTTTC 300
Oy 322 CTGAACCCACCTCAAGAGGTAATATTGAGGAATCACTTGGCAGATTGTTTAA 381
Db 301 CTGAACCCACCTCAAGAGGTAATATTGAGGAATCACTTGGCAGATTGTTTAA 360
Oy 382 TGCTGAGAACTGCTGCTCAAAATCAAGCAAACTAATGTTGCTCAAGGCTTGG 441
Db 361 TGCTGAGAACTGCTGCTCAAAATCAAGCAAACTAATGTTGCTCAAGGCTTGG 420
Oy 442 TCCTGAGAACTGAGAGAGAAATGCTCAAGATGCTCGGCTTCTCAAGCAGAG 501
Db 421 TCCTGAGAACTGAGAGAGAAATGCTCAAGATGCTCGGCTTCTCAAGCAGAG 480
Oy 502 CCTGCGGCTGCGAGGTTGTTATGACGGAAGTCTGGAATGAATGATATGTA 561
Db 481 CCTGCGGCTGCGAGGTTGTTATGACGGAAGTCTGGAATGAATGATATGTA 540
Oy 562 AGCTGATAGGATGTTGTTATGACGTTCTGACGTTCTTGTGAGCTTACACTTGT 621
Db 541 AGCTGATAGGATGTTGTTATGACGTTCTGACGTTCTTGTGAGCTTACACTTGT 600
Oy 622 TTAAGCAGG 630
Db 601 TTAAGCAGG 609

```

```

RESULT 13
LOCUS BE889624 863 bp mRNA linear EST 20-OCT-2000
DEFINITION 601512645F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3914194 5',
ACCESSION BE889624
VERSION BE889624.1 GI:10347134
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 863)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:

```









cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>  
Contact: (Dickson, Mark) mcd@paxil.stanford.edu  
Dickson, M., Schmitz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. M.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/ULNL at: <http://image.lnl.gov>  
Series: IRAC Plate: 14 Row: b Column: 7  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Hexamer frequency ORF  
analysis, Genomescan gene prediction.

FEATURES  
source location/Qualifiers  
1..2619

CDS  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="MGC:9960 IMAGE:3877854"  
/tissue\_type="Lung, large cell carcinoma"  
/clone\_id="NIH\_MGC\_68"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
216..797  
/codon\_start=1  
/product="Similar to RIKEN cDNA 1700037B15 gene"  
/protein\_id="AAH13592.1"  
/db\_xref="gi:15488932"  
/translation="MVATGSSSKNPASISLDDGHPISLSDPDYMDYVPEPNL  
NEVFEESTCONLYAKMLLENLSKSKYKLVKFKTORLAQDYLRLSSTPERG  
LRGCMVHNLIEIENVCCKLDRIKVDSSVLPFEELTLVFKDENCWSITFRDFESRGFR  
SSGRRLLILSSGRVLRKRLYSLGITVLEGS"

BASE COUNT 727 a 513 c 563 g 816 t

Query Match 98.8% Score 2579.8; DB 9; Length 2619;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 2590; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

1 AGCGCGCAGAGTGCGCGGAGAGGAGTACAGAGGTCGCCCTGCTGCGGCTGCA 60  
13 AGCGCGCAGAGTGCGCGGAGAGGAGTACAGAGGTCGCCCTGCTGCGGCTGCA 72  
61 GGTACCCGAGTCCCGAGAGAGGAGCTCTAAGAAAGCCCTGCTGTTACCGGCG 120  
73 GGTACCCGAGTCCCGAGAGAGGAGCTCTAAGAAAGCCCTGCTGTTACCGGCG 132  
121 AGCGAGCGGAGAGCGCCCGGGAAGCTCTGCGAGCGCTCAGAAAGGCGCTGCGCTCG 180  
133 AGCGAGCGGAGAGCGCCCGGGAAGCTCTGCGAGCGCTCAGAAAGGCGCTGCGCTCG 192  
181 CGAAGAAACAGAGCGCTTGACATGTTGCAACTGCGAGTTTGAGCAGCAAGACCGCG 240  
193 CGAAGAAACAGAGCGCTTGACATGTTGCAACTGCGAGTTTGAGCAGCAAGACCGCG 252  
241 CCAGCATTTCAGAATGCTGACATGCTATCACCCAGAGAGCCCTGCTAAGTATTTTG 300  
253 CCAGCATTTCAGAATGCTGACATGCTATCACCCAGAGAGCCCTGCTAAGTATTTTG 312  
301 ACTAGTGAGTATGTTGTTCTCTGAACCACTCAACGAGGATATTTTGGAGATCAA 360  
313 ACTAGTGAGTATGTTGTTCTCTGAACCACTCAACGAGGATATTTTGGAGATCAA 372  
361 CTTCGCAGAAATTTGGTTAAATGCTGAGAACTGTCTGCCAAATCAAGCAAACTAAC 420  
373 CTTCGCAGAAATTTGGTTAAATGCTGAGAACTGTCTGCCAAATCAAGCAAACTAAC 432  
421 TTGCTGCTCAAAAGTCTTGTCTCTGAGAACTGACGAGAAATTTGCTCAAGATGCC 480  
433 TTGCTGCTCAAAAGTCTTGTCTCTGAGAACTGACGAGAAATTTGCTCAAGATGCC 492  
481 TGGCGCTTCTCAAGAGAGCGCTGCGGCTTGCGAGGTGTTGATGACGTAAGACTTGG 540

493 TCGGCGCTTCTCAAGAGAGCGCTGCGGCTTGCGAGGTGTTGATGACGTAAGACTTGG 552  
541 AAATGAAATGTAATGTAAAGAGTGAATGATGTTGATGATTCGACGCTGACCTTA 600  
553 AAATGAAATGTAATGTAAAGAGTGAATGATGATGTTGATGATTCGACGCTGACCTTA 612  
601 CTTTGGACCTTACACTTGTGTTTAAAGAGAGAACTGCTCATGATGATGATGATGATG 660  
613 CTTTGGACCTTACACTTGTGTTTAAAGAGAGAACTGCTCATGATGATGATGATGATG 672  
661 TTTTCTTTAGTAGAGTGGCTTCTCTGCTGTTTAAAGAGAACTGCTCATGATGATG 720  
673 TTTTCTTTAGTAGAGTGGCTTCTCTGCTGTTTAAAGAGAACTGCTCATGATGATG 732  
721 GATTTGCACTGTTTAAAGAAACCTTCTCAGTATGGAACACAGATGATGATGATG 780  
733 GATTTGCACTGTTTAAAGAAACCTTCTCAGTATGGAACACAGATGATGATGATG 792  
781 CCTAAAGAGGAAATATATAAGATATTTTCATGATGATGATGATGATGATGATG 840  
793 CCTAAAGAGGAAATATATAAGATATTTTCATGATGATGATGATGATGATGATG 852  
841 GTTCACTTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
853 GTTCACTTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 912  
901 TCTCTTTTCTTTCTGTTTACATTTTACATTTTACATTTTACATTTTACATTTT 960  
913 TCTCTTTTCTTTCTGTTTACATTTTACATTTTACATTTTACATTTTACATTTT 972  
961 CTTTGGAGAGTATTAATAATAGTATGATGATGATGATGATGATGATGATGATG 1020  
973 CTTTGGAGAGTATTAATAATAGTATGATGATGATGATGATGATGATGATGATG 1032  
1021 AATGAATGGAGTGAATAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1080  
1033 AATGAATGGAGTGAATAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1092  
1081 TCCACAAATTTGTTTCAAGTATGATGATGATGATGATGATGATGATGATGATG 1140  
1093 TCCACAAATTTGTTTCAAGTATGATGATGATGATGATGATGATGATGATGATG 1152  
1141 TGATTTTCAATATGCTTTTCTGTAAGTTTCAAGGATTTTCAAGGATTTTCAAGG 1200  
1153 TGATTTTCAATATGCTTTTCTGTAAGTTTCAAGGATTTTCAAGGATTTTCAAGG 1212  
1201 ATACTAAATGATATTTAGCTTATCAAACTTCAAAAGGCGCACACAGTCTGACT 1260  
1213 ATACTAAATGATATTTAGCTTATCAAACTTCAAAAGGCGCACACAGTCTGACT 1272  
1261 CAATTTTGAAGGTTTAAAGACAGAGAGGATGATGATGATGATGATGATGATG 1320  
1273 CAATTTTGAAGGTTTAAAGACAGAGAGGATGATGATGATGATGATGATGATG 1332  
1321 TAACCTTAGTAAAGTAAATAGAAATATTTACGGCAATTAAGGAAGCTTCACTGTT 1380  
1333 TAACCTTAGTAAAGTAAATAGAAATATTTACGGCAATTAAGGAAGCTTCACTGTT 1392  
1381 CTTGCTGACTTCTCAGCTATGTTTAAAGAGCAACAAAGCTTATGATGATGATG 1440  
1393 CTTGCTGACTTCTCAGCTATGTTTAAAGAGCAACAAAGCTTATGATGATGATG 1452  
1441 AGTAACACAGGTTGTTTAAAGTATGATGATGATGATGATGATGATGATGATG 1500  
1453 AGTAACACAGGTTGTTTAAAGTATGATGATGATGATGATGATGATGATGATG 1512  
1501 TGAATACGCTCTTAACTAGAAATATGCAAAATATTAACATGATATTTCTGCTAGT 1560  
1513 TGAATACGCTCTTAACTAGAAATATGCAAAATATTAACATGATATTTCTGCTAGT 1572  
1561 TTATACGCTCTTAACTAGATGATGATGATGATGATGATGATGATGATGATGATG 1620

Db	1573	TGATTACTCTCCATGTAGTGGAATGATGAAATACCTGCGAGGATAGAGAAATTAAGTCTGTAG	1632
Qy	1621	ATCTAGTCCCTCTGTGATCTACTAGTAATCTACTCTTTTCAAGTAATACATGTGCTTA	1680
Db	1633	ATCTAGTCCCTCTGTGATCTACTAGTAATCTACTCTTTTCAAGTAATACATGTGCTTA	1692
Qy	1681	ATCTCAGATGAAACATTTACATGAGCAGTGTATCTCATCTCTGGGCTTTTGTGGAAAT	1740
Db	1693	ATCTCAGATGAAACATTTACATGAGCAGTGTATCTCATCTCTGGGCTTTTGTGGAAAT	1752
Qy	1741	TGAAGTATCTCTCTTAACCCCAATTTGTCAAGGGTAGTACTGTATACACACTTTTGA	1800
Db	1753	TGAAGTATCTCTCTTAACCCCAATTTGTCAAGGGTAGTACTGTATACACACTTTTGA	1812
Qy	1801	TTTATTGAAACGGGTCATTTTACGACAGCTGCTATGGCTATGAGAGATATGGTTATAGTAC	1860
Db	1813	TTTATTGAAACGGGTCATTTTACGACAGCTGCTATGGCTATGAGAGATATGGTTATAGTAC	1872
Qy	1861	AGCCTAGAAATGAAACTCAACCGTCAGATACCATGATGACACCCAGATTTTTCACAC	1920
Db	1873	AGCCTAGAGATGAAACTCAACCGTCAGATACCATGATGACACCCAGATTTTTCACAC	1932
Qy	1921	TTTGATACCTCTCTCTAGGGAATATATAGGCGCTGATTTTCTCTTATTTCCAACTAAGT	1980
Db	1933	TTTGATACCTCTCTCTAGGGAATATATAGGCGCTGATTTTCTCTTATTTCCAACTAAGT	1992
Qy	1981	AGATCATATATCTCTTCTCTTTTATGTATGTATGAGAGATTTTGGCCGTCACACAAAG	2040
Db	1993	AGATCATATATCTCTTCTCTTTTATGTATGTATGAGAGATTTTGGCCGTCACACAAAG	2052
Qy	2041	TTTCAATTCAGCAGGCTTTCATATCTCTGCTGTGGGTCGTGATTAAGAGACTTATTCAGG	2100
Db	2053	TTTCAATTCAGCAGGCTTTCATATCTCTGCTGTGGGTCGTGATTAAGAGACTTATTCAGG	2112
Qy	2101	TTTCTGCGCCAGTATTTAGTCCACTTCAATGCTGGAGACGCGGTGGGACAAATGT	2160
Db	2113	TTTCTGCGCCAGTATTTAGTCCACTTCAATGCTGGAGACGCGGTGGGACAAATGT	2172
Qy	2161	ATTCATCTCTGCTTACTGAAAAACAGGTGTGATCTCTGTATGATATTAAGTGACCT	2220
Db	2173	ATTCATCTCTGCTTACTGAAAAACAGGTGTGATCTCTGTATGATATTAAGTGACCT	2232
Qy	2221	AAAAATGTCACGTTTCAATTTAGCAAGTGTCTTAACAACACTTCTTCAATGCTTGA	2280
Db	2233	AAAAATGTCACGTTTCAATTTAGCAAGTGTCTTAACAACACTTCTTCAATGCTTGA	2292
Qy	2281	AAGATACCTAACAAGCCATCTTTATAGAAATTTGGCCCAAGATTAATCATATTTTGTGCA	2340
Db	2293	AAGATACCTAACAAGCCATCTTTATAGAAATTTGGCCCAAGATTAATCATATTTTGTGCA	2352
Qy	2341	TGCGTATTTGTAAGCTCCAAAGTTTCACTGTGTTTCTGCGGCTGCTCTGAGTTGTACCC	2400
Db	2353	TGCGTATTTGTAAGCTCCAAAGTTTCACTGTGTTTCTGCGGCTGCTCTGAGTTGTACCC	2412
Qy	2401	ACTGACGTGGGAAGGCTTTTGGGCAATGAGATGAGAACTTTGTCTTTCCACATTAACA	2460
Db	2413	ACTGACGTGGGAAGGCTTTTGGGCAATGAGATGAGAACTTTGTCTTTCCACATTAACA	2472
Qy	2461	GTTATCTTTGACCTCTTGTGCTGTATGCTTACAAATGCTGATGAGCTTATGGAAGGCTG	2520
Db	2473	GTTATCTTTGACCTCTTGTGCTGTATGCTTACAAATGCTGATGAGCTTATGGAAGGCTG	2532
Qy	2521	TTTAATTTAATTTCTCTGTTAAAGAAATTAAGTTTGTCTATTTTGTACAAATTAACATT	2580
Db	2533	TTTAATTTAATTTCTCTGTTAAAGAAATTAAGTTTGTCTATTTTGTACAAATTAACATT	2592
Qy	2581	ATATATTTTAAAAAATTTAAAAAATTTAAAAAATTTAAAAAATTTAAAAAATTTAAAAAATTT	2607
Db	2593	ATATATTTTAAAAAATTTAAAAAATTTAAAAAATTTAAAAAATTTAAAAAATTTAAAAAATTT	2619

LOCUS	AP001961	70422 bp	DNA	linear	PRI 24-OCT-2000
DEFINITION	Homo sapiens genomic DNA, chromosome 4q22-q24, clone:2190A3, complete sequence.				
ACCESSION	AP001961				
VERSION	AP001961.2	GI:9929904			
KEYWORDS	HTG.				
SOURCE	Homo sapiens DNA, clone:2190A3.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Tsai,S.F.				
TITLE	Genome Sequencing of the Chromosome 4q Region Implicated in Human Hepatocellular Carcinoma Pathogenesis 1				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 70422)				
AUTHORS	Tsai,S.F.				
TITLE	Direct Submission				
JOURNAL	Submitted (28-APR-2000) to the DDBJ/EMBL/GenBank databases. Shih-Feng Tsai, National Yang-Ming University, Institute of Genetics, 155 Li-Rong St. Section 2, Peitou, Taipei, Taiwan 11221, Republic of China (E-mail: ympeusa1@ym.edu.tw, URL: http://genome.ym.edu.tw/, Tel:886-2-28267043, Fax:886-2-28264930)				
COMMENT	On Aug 26, 2000 this sequence version replaced gi:7678851. Quality: the expected Phred/Phrap calculated error rate (per 10 kb) is 0.0002; Estimated total number of errors is 0.0013.				
FEATURES					
Source	1..70422				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/chromosome="4"				
	/map="4q22-q24"				
	/clone="2190A3"				
BASE COUNT	22872 a 13751 c 13163 g 20636 t				
ORIGIN					
Query Match	87.5%; Score 2284.6; DB 9; Length 70422;				
Best Local Similarity	99.4%; Pred. No. 0;				
Matches 2293; Conservative	0; Mismatches 14; Indels 0; Gaps 0;				
Qy	286	TGCTAAGGATTTTGACTACTGCGGATTTATGTGTCTGTAACCCAACTCAACGAGCTAA	345		
Db	23632	TTCTTACAGATTTTGACTACTGCGGATTTATGTGTCTGTAACCCAACTCAACGAGCTAA	23691		
Qy	346	TATTTGAGGAATCAACTTGCAGCAATTTGGTTAAATAGCTGGAAGAACTGCTGTCCAAT	405		
Db	23632	TATTTGAGGAATCAACTTGCAGCAATTTGGTTAAATAGCTGGAAGAACTGCTGTCCAAT	23751		
Qy	406	CAAAAGCAAACTTAACCTGTGTTGCTCAAAAGGTCCTTGTCCCTGAGAAACTGACGACAGAA	465		
Db	23752	CAAAAGCAAACTTAACCTGTGTTGCTCAAAAGGTCCTTGTCCCTGAGAAACTGAC	23811		
Qy	466	TTGCTCAAGATGTCCTGTGGGCTTTCCTCAACGGAGCCCTGCGGCTTGCAGAGTTGTGTA	525		
Db	23812	TTGCTCAAGATGTCCTGTGGGCTTTCCTCAACGGAGCCCTGCGGCTTGCAGAGTTGTGTA	23871		
Qy	526	TGCACGTGAACCTTGGAATTTGAAATGTATGTAAAAAAGCTGGAATAGGATTTGTCTGATT	585		
Db	23872	TGCACGTGAACCTTGGAATTTGAAATGTATGTAAAAAAGCTGGAATAGGATTTGTCTGATT	23931		
Qy	586	CTAGCGTGTACCTACTTTTGTAGCTTACACTGTGTTTAAAGCAGAGAACTGCTCATGGA	645		
Db	23932	CTAGCGTGTACCTACTTTTGTAGCTTACACTGTGTTTAAAGCAGAGAACTGCTCATGGA	23991		
Qy	646	CTAGCTTCAGAGCACTTTTCTTATAGTAGAGGTGCTCTCTCTGTGTTTCAGAGCAACTC	705		
Db	23992	CTAGCTTCAGAGCACTTTTCTTATAGTAGAGGTGCTCTCTCTGTGTTTCAGAGCAACTC	24051		
Qy	706	TGATCTCAGCTCAGAGATTTTGACTGTGTAAGAAAAAATTTTACTACATGATTTGAGACA	765		
Db	24052	TGATCTCAGCTCAGAGATTTTGACTGTGTAAGAAAAAATTTTACTACATGATTTGAGACA	24111		

QY	766	CAGTGAATTGCAAGGGCTCTAAAAAGGAAATATTAAGAATATTTCAATGATGGGTAGT	825
Dd	24112	CAGTGAATTGCAAGGGCTCTAAAAAGGAAATATTAAGAATATTTCAATGATGGGTAGT	24171
QY	826	AAATCATATTCACGCTAGTCAGCTAAAGTCATATTTGATGTTGGCCCACTGCCCTAAATFAG	885
Dd	24112	AAATCATATTCACGCTAGTCAGCTAAAGTCATATTTGATGTTGGCCCACTGCCCTAAATFAG	24231
QY	886	AAACCCCAAAATAGTCTCTTTTCTTTCTGTGTTCACATTCATAGCAATGCACTAAC	945
Dd	24232	AAACCCCAAAATAGTCTCTTTTCTTTCTGTGTTCACATTCATAGCAATGCACTAAC	24291
QY	946	AGGCGATTTTCTGGCCCTTGGAGAAAGTATCAAAATATAGTACATTTTCGCAATAGT	1005
Dd	24292	AGGCGATTTTCTGGCCCTTGGAGAAAGTATCAAAATATAGTACATTTTCGCAATAGT	24351
QY	1006	CCCATTTTGTATCAGAAATATGAATGGATGGATAGGTAAAGTAAAGTAAAGCCATTGTG	1065
Dd	24352	CCCATTTTGTATCAGAAATATGAATGGATGGATAGGTAAAGTAAAGTAAAGCCATTGTG	24411
QY	1066	AGTTTTCATTTTGATTTCCAAATTTGGTTTCAGTAGGCGTTGGTATAGCTATATTAAC	1125
Dd	24412	AGTTTTCATTTTGATTTCCAAATTTGGTTTCAGTAGGCGTTGGTATAGCTATATTAAC	24471
QY	1126	CAGATTTGGCTATTTTGTGATTTTCATATGGCTTTTTTTTCTGTAACTTTTCAGAGATTTT	1185
Dd	24472	CAGATTTGGCTATTTTGTGATTTTCATATGGCTTTTTTTTCTGTAACTTTTCAGAGATTTT	24531
QY	1186	TTTAAATCAGAGATCATACTAATATATATTAAGCTATCAAACTTCCAAAAGCCACAC	1245
Dd	24532	TTTAAATCAGAGATCATACTAATATATATTAAGCTATCAAACTTCCAAAAGCCACAC	24591
QY	1246	CACCAAGTTCCTGACTCAAAATTTTGAAGGGTTTTTTCAGACAGAGGGTAGATTAAAGTGTG	1305
Dd	24592	CACCAAGTTCCTGACTCAAAATTTTGAAGGGTTTTTTCAGACAGAGGGTAGATTAAAGTGTG	24651
QY	1306	AGTTTAAATTAAGGCTTAACCCCTAGGTAAAGATGAAGTAATATTACGGCAATATATGA	1365
Dd	24652	AGTTTAAATTAAGGCTTAACCCCTAGGTAAAGATGAAGTAATATTACGGCAATATATGA	24711
QY	1366	ACTGCTTCACTGTTTCTTGGTACTTCCCTCACTCAATATGTTTAAAGAGGCAACAAAAGC	1425
Dd	24712	ACTGCTTCACTGTTTCTTGGTACTTCCCTCACTCAATATGTTTAAAGAGGCAACAAAAGC	24771
QY	1426	TTATGGTGCCATTTCACTAACCACGGGTGTTGTTTAGATGCTTTATTAAGCTCAGTTTCC	1485
Dd	24772	TTATGGTGCCATTTCACTAACCACGGGTGTTGTTTAGATGCTTTATTAAGCTCAGTTTCC	24831
QY	1486	CTTGTGTTCAATAGTGGATFACGTGCTTTAACTGAAAAATGCAAAATATTTGAAGTGT	1545
Dd	24832	CTTGTGTTCAATAGTGGATFACGTGCTTTAACTGAAAAATGCAAAATATTTGAAGTGT	24891
QY	1546	ATTTCGTGTGTAGTTTATTAATCTCTTCCATTTAGAGTAAGATGAATFACCTGTAGAGATAG	1605
Dd	24892	ATTTCGTGTGTAGTTTATTAATCTCTTCCATTTAGAGTAAGATGAATFACCTGTAGAGATAG	24951
QY	1606	GAATATGAGTTCTGAGATCTAGTCCCTCTCTGTGATTCACCTAGTAATCTACTCTTTTCAG	1665
Dd	24952	GAATATGAGTTCTGAGATCTAGTCCCTCTCTGTGATTCACCTAGTAATCTACTCTTTTCAG	25011
QY	1666	TATTAATCATGTGCTTAATCTCAGATGACCATTTTCACCATGCGAGTGTATATCTCATCTCTG	1725
Dd	25012	TATTAATCATGTGCTTAATCTCAGATGACCATTTTCACCATGCGAGTGTATATCTCATCTCTG	25071
QY	1726	GGCTTTTCTGGGAATTTGAAGTATCTCTCTTAACCCCAATTTGCAAGGGTAACTAGCTGTGA	1785
Dd	25072	GGCTTTTCTGGGAATTTGAAGTATCTCTCTTAACCCCAATTTGCAAGGGTAACTAGCTGTGA	25131
QY	1786	TACTACCACTTTGAATATTAATGAAGGGGCATTTTAACGAAGTCTGATTTGGCTATGGAGA	1845
Dd	25132	TACTACCACTTTGAATATTAATGAAGGGGCATTTTAACGAAGTCTGATTTGGCTATGGAGA	25191
QY	1846	TATGTGTTTATAGTACAGCCTTAGAAGATGAAGTAAGTCAACGCTCGAGATTAACCATGATCAGCC	1905

Db	25192	1107 bp	linear	ROD 09-DEC-2001
QY	1906	CAGATTTTTCACCCCTGGATACCTGCATAGGGAATTAATAAGGCGCTGATTTTTTTC	1965	
Db	25252	CAGATTTTTTTCACCCCTGGATACCTGCATAGGGAATTAATAAGGCGCTGATTTTTTTC	25311	
QY	1966	TTATTTCCAACTAAGTAGATCATTTATCTCTTCTTTTATGTTAATGAGAAATTTAGC	2025	
Db	25312	TTATTTCCAACTAAGTAGATCATTTATCTCTTCTTTTATGTTAATGAGAAATTTAGC	25371	
QY	2026	CTCCACTCACAATGTCATATTCAGCAAGGCTTTATATCTTGGTGGGCTGGGATA	2085	
Db	25372	CTCCACTCACAATGTCATATTCAGCAAGGCTTTATATCTTGGTGGGCTGGGATA	25431	
QY	2086	AGGAGCTTATTCAGGTTTCTGCGCTAGCTATTTAGCTCCACTTCACATGCTGGAGACTGG	2145	
Db	25432	AGGAGCTTATTCAGGTTTCTGCGCTAGCTATTTAGCTCCACTTCACATGCTGGAGACTGG	25491	
QY	2146	CGTAGGACAGATGTTATTCATCTCTGCTGTTACTGAAAAACAGGTGATCCTGTTAGTGA	2205	
Db	25492	CGTAGGACAGATGTTATTCATCTCTGCTGTTACTGAAAAACAGGTGATCCTGTTAGTGA	25551	
QY	2206	TACTATTAAGTGACCTAAAAATGTCACGTTCCAATTAGCAAGGTTCCTAACAACTAACT	2265	
Db	25552	TACTATTAAGTGACCTAAAAATGTCACGTTCCAATTAGCAAGGTTCCTAACAACTAACT	25611	
QY	2266	CTTCAAAATGCTTGGAAAAAGATACACAAAGCCAACTTTTATAGAAATTTGGGCCAAGATAAT	2325	
Db	25612	CTTCAAAATGCTTGGAAAAAGATACACAAAGCCAACTTTTATAGAAATTTGGGCCAAGATAAT	25671	
QY	2326	CTATGTTGTTTGGCATGCTATTTGTTAGCTCCAAAGGTTCTACTGTGTTCTGCCGCTGT	2385	
Db	25672	CTATGTTGTTTGGCATGCTATTTGTTAGCTCCAAAGGTTCTACTGTGTTCTGCCGCTGT	25731	
QY	2386	CGTGAGTTGTACACACACACACACAGGCGCTTGGGCAAGATGTTGAACCTGTTGTC	2445	
Db	25732	CGTGAGTTGTACACACACACACAGGCGCTTGGGCAAGATGTTGAACCTGTTGTC	25791	
QY	2446	CTTTTCCCACTAACAGTTATCTTTGACCTCTCTGCTGTTATGCTTACAAAAATGCTGATG	2505	
Db	25792	CTTTTCCCACTAACAGTTATCTTTGACCTCTCTGCTGTTATGCTTACAAAAATGCTGATG	25851	
QY	2506	GCTTATGAGAGCGCTGTTAAATTAATATTCCTGTTAAAGAAATTTAAAGTTGTCTATTTT	2565	
Db	25852	GCTTATGAGAGCGCTGTTAAATTAATATTCCTGTTAAAGAAATTTAAAGTTGTCTATTTT	25911	
QY	2566	TGACATAATAAACATTATATATTTTTTAA	2592	
Db	25912	TGACATAATAAACATTATATATTTTTTAA	25938	
RESULT 3	AF327511	1107 bp	mRNA	linear
LOCUS	AF327511			
DEFINITION	Rattus norvegicus SMHS1 (Smhs1) mRNA, complete cds.			
ACCESSION	AF327511			
VERSION	AF327511.1			
KEYWORDS	GI:17432244			
SOURCE	Norway rat.			
ORGANISM	Rattus norvegicus			
REFERENCE	1 (bases 1 to 1107)			
AUTHORS	Cros,N., Tkatchenko,A.V., Leclerc,L., Leger,J.J., Marini,J.-F., and			
TITLE	Dechesne,C.A.			
JOURNAL	Gene expression alterations revealed by suppression subtractive			
REFERENCE	hybridization in rat soleus muscle disuse atrophy			
AUTHORS	unpublished			
TITLE	2 (bases 1 to 1107)			
REFERENCE	Leclerc,L., Pisanl,D.F., Cros,N. and Dechesne,C.A.			
AUTHORS	Direct Submission			
TITLE				



JOURNAL Submitted (11-DEC-2000) CNRS, faculte des sciences, UMR 6548, Parc  
Valrose, Nice 06108, France  
FEATURES Location/Qualifiers  
source 1..1107  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="03A12"  
/issue\_type="hindlimb skeletal muscle"  
1..1107  
/gene="Smhsl"  
439..1020  
/gene="Smhsl"  
/codon\_start=1  
/product="Smhsl"  
/protein\_id="AAU39012.1"  
/db\_xref="GI:17432245"  
/translation="MVATGSLSSKNPASISELLDGGSHPSLSDPDMYVPEPRL  
NEVFEEETCONLVMKLENCISKOTKLGCSVLPKLTORLADYLRSLSTPECG  
LRGCVHVNLEIENVCCKLDRIYCDASVPEFLTLVFKQESCSWTSKDFPSGRF  
SSGLRTRLSSFRVKKVLSLIGTVIEEC"

BASE COUNT 344 a 207 c 260 g 296 t  
ORIGIN

Query Match 20.1% Score 523.6; DB 10; Length 1107;  
Best Local Similarity 84.9% Pred. No. 4.2e-115;  
Matches 566; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 161 GGAAGGCCGTTGCCGCGAAGAACAGAGCCGTTGACATGCTGCACTGCAG 220  
DB 396 GAAAGATCCGTTCTGCTCCGCGAGGAAACAGAGCCGTTGACATGCTGCAAGGGCAG 455  
QY 221 TTTCAGCAGCAGAACCCGCGCAGCATTTTCAAGATTCTGAGCTGCTATCCACCA 280  
DB 456 TTTCAGCAGCAGAACCCGCGCAGCATTTTCAAGATTCTGAGCTGCTATCCACCTGG 515  
QY 281 GAGCTGCTAAGTATTTGACTACCTAGGATTTGTTGTTCTGACCCCAACCTCAACGA 340  
DB 516 GAGCTGCTAAGTATTTGACTACCTAGGATTTGTTGTTCTGACCCCAACCTCAACGA 575  
QY 341 GGTAAATATTTGAGGAACTCACTTCCAGAAATTTGGTTAAATGCTGGAAGATGCTGTC 400  
DB 576 GGTGTTGTTGAGAGACAAACATGCGAAATTTGGTTAAATGTTGGAAGATGCTGTC 635  
QY 401 CAATCAAGCAAACTAACTGCTGCTCAAGAGTCTTCTGCTGAGAACTGACGCA 460  
DB 636 CAAGTCAAGCAAACTAACTGCTGCTGAGTCTGCTGAGTCTGAGAACTGACGCA 695  
QY 461 GAGAATGCTCAAGATGCTGCTGCTGCTCAAGAGAGCCGCGGCTTGCAGAGTTG 520  
DB 696 GAGAATGCTCAAGATGCTGCTGCTGCTCAAGAGAGCCGCGGCTTGCAGAGTTG 755  
QY 521 TGTATGACGCTGAACTTGAATTTGAATTTGAATTTGAATTTGAATTTGAATTTGA 580  
DB 756 TGTATGACGCTGAACTTGAATTTGAATTTGAATTTGAATTTGAATTTGAATTTGA 815  
QY 581 TGTATGACGCTGAACTTGAATTTGAATTTGAATTTGAATTTGAATTTGAATTTGA 640  
DB 816 TGTATGACGCTGAACTTGAATTTGAATTTGAATTTGAATTTGAATTTGAATTTGA 875  
QY 641 ATGAGCTAGCTGAGGAACTTTTCTTATAGAGTGGCTTCTGCTGCTGCTGCTGCTG 700  
DB 876 CTGAGCAGCCTCAAGAGCTTTCTTCTTACGGAGTGGCTTCTGCTGCTGCTGCTGCTG 935  
QY 701 AACTCTGATCTCAGCTCAGGATTTTCTTATAGAGTGGCTTCTGCTGCTGCTGCTGCTG 760  
DB 936 AACTCTGATCTCAGCTCAGGATTTTCTTATAGAGTGGCTTCTGCTGCTGCTGCTGCTG 995  
QY 761 AACCAAGCTGATGAGAGTCTTAAAGGAAATTTATTAAGATTTATTTGATGATTTG 820  
DB 996 AACCAAGCTGATGAGAGTCTTAAAGGAAATTTATTAAGATTTATTTGATGATTTG 1055  
QY 821 GTAGTAAACCTATTTCACTAGTACGCTAAA 850

DB 1056 CTAATAAACCTCTGAGCTATCCAGCAAA 1085

RESULT 4  
AF353325 667 bp mRNA linear ROD 06-DEC-2001  
LOCUS Mus musculus Rtp801-like protein mRNA, complete cds.  
DEFINITION  
AF353325  
VERSION  
AF353325.1 GI:17385766  
KEYWORDS  
SOURCE  
ORGANISM  
house mouse.  
Mus musculus

REFERENCE  
1 (bases 1 to 667)  
Shoshani,T., Faerman,A., Met,I., Zelin,E., Tenne,T., Gorodin,S.,  
Moshe,Y., Elbaz,S., Budanov,A., Chajut,A., Kalinski,H., Kamer,I.,  
Rosen,A., Mor,O., Keshet,E., Leshkowitz,D., Elnat,P., Skallier,R.  
and Feinstein,E.,  
Identification of a novel HIF-1 responsive gene Rtp801 involved in  
apoptosis  
Mol. Cell. Biol. (2002) In press

JOURNAL  
REFERENCE  
2 (bases 1 to 667)  
Shoshani,T., Faerman,A., Met,I., Zelin,E., Tenne,T., Gorodin,S.,  
Moshe,Y., Elbaz,S., Budanov,A., Chajut,A., Kalinski,H., Kamer,I.,  
Rosen,A., Mor,O., Keshet,E., Leshkowitz,D., Elnat,P., Skallier,R.  
and Feinstein,E.,  
Direct Submission  
Submitted (10-JAN-2001) Research Division, QBI Enterprises Ltd.,  
P.O. Box 4071, Ness-Ziona 70400, Israel

TITLE  
JOURNAL  
P.O. Box 4071, Ness-Ziona 70400, Israel  
FEATURES  
source  
Location/Qualifiers  
1..667  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
64..645  
/note="Rtp801L"  
/codon\_start=1  
/product="Rtp801-like protein"  
/protein\_id="AAU38425.1"  
/db\_xref="GI:17385767"  
/translation="MVATGSLSSKNPASISELLDGGSHPSLSDPDMYVPEPRL  
NEVFEEETCONLVMKLENCISKOTKLGCSVLPKLTORLADYLRSLSTPECG  
LRGCVHVNLEIENVCCKLDRIYCDASVPEFLTLVFKQESCSWTSKDFPSGRF  
SSGLRTRLSSFRVKKVLSLIGTVIEEC"

CDS

BASE COUNT 170 a 149 c 184 g 164 t  
ORIGIN

Query Match 20.0% Score 522.4; DB 10; Length 667;  
Best Local Similarity 88.2% Pred. No. 7.6e-115;  
Matches 566; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 161 GGAAGGCCGTTGCCGCGAAGAACAGAGCCGTTGACATGCTGCACTGCAG 220  
DB 21 GAAAGATCCGTTCTGCTCCGCGAGGAAACAGAGCCGTTGACATGCTGCAAGGGCAG 80  
QY 221 TTTCAGCAGCAGAACCCGCGCAGCATTTTCAAGATTCTGAGCTGCTATCCACCA 280  
DB 221 TTTCAGCAGCAGAACCCGCGCAGCATTTTCAAGATTCTGAGCTGCTATCCACCA 280  
QY 281 GAGCTGCTAAGTATTTGACTACCTAGGATTTGTTGTTCTGACCCCAACCTCAACGA 340  
DB 81 TTTCAGCAGCAGAACCCGCGCAGCATTTTCAAGATTCTGAGCTGCTATCCACCTGG 140  
QY 341 GAGCTGCTAAGTATTTGACTACCTAGGATTTGTTGTTCTGACCCCAACCTCAACGA 340  
DB 141 GAGCTGCTAAGTATTTGACTACCTAGGATTTGTTGTTCTGACCCCAACCTCAACGA 200  
QY 341 GGTAAATATTTGAGGAACTCACTTCCAGAAATTTGGTTAAATGCTGGAAGATGCTGTC 400  
DB 201 GGTAAATATTTGAGGAACTCACTTCCAGAAATTTGGTTAAATGCTGGAAGATGCTGTC 260  
QY 401 CAATCAAGCAAACTAACTGCTGCTCAAGAGTCTTCTGCTGAGAACTGACGCA 460  
DB 261 CAATCAAGCAAACTAACTGCTGCTCAAGAGTCTTCTGCTGAGAACTGACGCA 320  
QY 461 GAGAATGCTCAAGATGCTGCTGCTGCTCAAGAGAGCCGCGGCTTGCAGAGTTG 520

|||||  
Db 321 GAGAAATTCCTCAAGATGCTGCGGCTCTCGTCCACGAGCCCTCGGCGCTTCGGGGTTG 380  
Qy 521 TGTATGACAGCTGAACTGGAAATGTATGTAAAAAGCTGGATAGATTGTGTG 580  
Db 381 TGTATGACAGCTGAACTGGAAATGTATGTAAAAAGCTGGATAGATTGTGTG 440  
Qy 581 TGAATTCAGCTGCTGACCTACTTGTGACCTTACACTTGTGTTAAGCAGAGAACCTGTC 640  
Db 441 TGAATTCAGCTGCTGACCTACTTGTGACCTTACACTTGTGTTAAGCAGAGAACCTGTC 500  
Qy 641 ATGACCTAGCTTCAAGGAGCTTTTCTTTAGTAGAGGTGCTTCCTCTGCTTTCAGAG 700  
Db 501 GTGACACAGCTTCAAGGAGCTTCTTCTTTAGCAGAGGTGCTTCCTCAGGCTTAAAG 560  
Qy 701 AACTGTATCTGACAGCTGAGATTGCGACTGTGTAAGAAAAAAGCTTCTCAGATTGG 760  
Db 561 AACTGTATCTGACAGCTGAGATTGCGACTGTGTAAGAAAAAAGCTTCTCAGATTGG 620  
Qy 761 AACACAGCTGATTGAAGGCTCTTAAAGGAAATATATTAAG 804  
Db 621 AACGACAGCTATTGAAGAGTGTCTAAGAGGAAACAACTTAAAG 664  
RESULT 5  
AX268802 867 bp DNA linear PAT 29-OCT-2001  
LOCUS AX268802  
DEFINITION Sequence 3 from Patent WO0174901.  
ACCESSION AX268802  
VERSION AX268802.1 GI:16541861  
KEYWORDS Norway rat.  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (sites)  
AUTHORS Stanton, L.W. and White, R.T.  
TITLE Secreted factors  
JOURNAL Patent: WO 0174901-A 3 11-OCT-2001;  
Scienc. Inc. (US)  
FEATURES  
source location/Qualifiers  
1..867  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
BASE COUNT 204 a 202 c 245 g 199 t 17 others  
ORIGIN  
Query Match 19.6%; Score 511; DB 6; Length 867;  
Best Local Similarity 85.0%; Pred. No. 4,3e-112;  
Matches 571; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Qy 161 GAAAGCCGCTTGGCGCTCGGAAAGAAAGAGCGCTTGAACATGTTGCAACTGGCAG 220  
Db 172 GAAAGATCCGTTCTGCTCGCGAGGAAACAGAGCGCTTGAACATGTTGCAAGCGCAG 231  
Qy 221 TTGAGCAGCAGAAACCCGCGCAGCATTTGCAATTTGCTGCACTGCTTACACCAGA 280  
Db 232 TTGAGCAGTAAAGACGCGCAGCATTTGCAAGTTGCTGAGCGGTGCTCACCCTGG 291  
Qy 281 GAGCTGCTAGTATTTTGAAGTATGTTGCTGCAAGCTTACACCAGA 340  
Db 292 GAGCTGCTAGTATTTTGAAGTATGTTGCTGCAAGCTTACACCAGA 351  
Qy 341 GGTATATTTTGAAGTATTTTGAAGTATTTTGAAGTATTTTGAAGTATTTTGAAGTAT 400  
Db 352 GGTATATTTTGAAGTATTTTGAAGTATTTTGAAGTATTTTGAAGTATTTTGAAGTAT 411  
Qy 401 CAAATCAAGCAAACTAATTTGTTGCTCAAGGCTTCTGCTGCAAGCAAACTAATTTG 460  
Db 412 CAAATCAAGCAAACTAATTTGTTGCTCAAGGCTTCTGCTGCAAGCAAACTAATTTG 471  
Qy 461 GAGATTTGCTCAAGTATTTGCTGCGGCTTCTCAAGGAGCCCTGCGGCTTGCAGGTTG 520

|||||  
Db 472 GAGAAATTCCTCAAGATGCTGCGGCTCTCATTCACAGAGCCCTCGGCGCTTCGGGGCTG 531  
Qy 521 TGTATGACAGCTGAACTGGAAATGTATGTAAAAAGCTGGATAGATTGTGTG 580  
Db 532 TGTATGACAGCTGAACTGGAAATGTATGTAAAAAGCTGGATAGATTGTGTG 591  
Qy 581 TGAATTCAGCTGCTGACCTACTTGTGACCTTACACTTGTGTTAAGCAGAGAACCTGTC 640  
Db 592 TGAATTCAGCTGCTGACCTACTTGTGACCTTACACTTGTGTTAAGCAGAGAACCTGTC 651  
Qy 641 ATGACCTAGCTTCAAGGAGCTTTTCTTTAGTAGAGGTGCTTCCTCTGCTTTCAGAG 700  
Db 652 GTGACACAGCTTCAAGGAGCTTCTTCTTTAGCAGAGGTGCTTCCTCAGGCTTAAAG 711  
Qy 701 AACTGTATCTGACAGCTGAGATTGCGACTGTGTAAGAAAAAAGCTTCTCAGATTGG 760  
Db 712 AACTGTATCTGACAGCTGAGATTGCGACTGTGTAAGAAAAAAGCTTCTCAGATTGG 771  
Qy 761 AACACAGCTGATTGAAGGCTCTTAAAGGAAATATATTAAGATTATTCATGATTGG 820  
Db 772 AACGACAGCTATTGAAGAGTGTCTAAGAGGAAACAACTTAAAGGTCTCATTAGATTGG 831  
Qy 821 GTAGTAAACTA 832  
Db 832 CTAAACAAAAANA 843  
RESULT 6  
AF327512 639 bp mRNA linear ROD 09-DEC-2001  
LOCUS AF327512  
DEFINITION Mus musculus SMH51 (Smh51) mRNA, complete cds.  
ACCESSION AF327512  
VERSION AF327512.1 GI:17432246  
KEYWORDS house mouse.  
SOURCE Mus musculus  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 639)  
AUTHORS Cros, N., Tkatchenko, A.V., Leclerc, L., Leger, J.J., Marini, J.-F. and  
Dechesne, C.A.  
TITLE Gene expression alterations revealed by suppression subtractive  
hybridization in rat soleus muscle atrophy  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 639)  
REFERENCE Leclerc, L., Pisanil, D.F., Cros, N. and Dechesne, C.A.  
JOURNAL Direct Submission  
TITLES Submitted (11-DEC-2000) CNRS, Faculte des Sciences, UMR 6548, Parc  
Valrose, Nice 06108, France  
FEATURES  
source location/Qualifiers  
1..639  
/organism="Mus musculus"  
/strain="C57BL/10ScSn"  
/db\_xref="taxon:10090"  
/clone="03A12"  
/tissue\_type="hindlimb skeletal muscle"  
1..639  
/gene="Smh51"  
19..600  
/gene="Smh51"  
/codon\_start=1  
/product="SMH51"  
/protein\_id="AA139013.1"  
/db\_xref="GI:17432247"  
/translation="MVATGSLSSKNPASISLLDGGYHRGMSLDFQWYVYVEPNL  
NEVVEFTTQNLVKNLENCISRSKQTLGCSKLVPEKLTQRTAODVLRLSSTPCG  
LRGCVMHVNEIEWCKKLDRIYCDASVPPFELTLVFKQSCPTSLKDFFSRGR  
SSGLKRTLLISSGFRLVKKKLYSLIGTVIEEC"  
BASE COUNT 166 a 136 c 173 g 164 t  
ORIGIN  
Query Match 19.5%; Score 509; DB 10; Length 639;



Db 225 CACGTGTTCTGCGCCGTCTGCTGAGTTGTCACCACTGACTGGGCAAGGCTTCTTGGGC 166  
Qy 2426 ATGATGATGAGAACTGTTGCTTTTCCACATACAGATTATCTTGACATCTTGCGCTGT 2485  
Db 165 ATGATGATGAGAACTGTTGCTTTTCCACATACAGATTATCTTGACATCTTGCGCTGT 106  
Qy 2486 ATGCTTACAAAATGCTGATGCTTATGGAAGGCTGTTAAATATATCTTGTTAAAGCA 2545  
Db 105 ATGCTTACAAAATGCTGATGCTTATGGAAGGCTGTTAAATATATCTTGTTAAAGCA 46  
Qy 2546 AATTAAGTTGCTGCTATTTTGCACATTAACATTAATATATTTT 2590  
Db 45 AATTAAGTTGCTGCTATTTTGCACATTAACATTAATATATTTT 1  
RESULT 9  
LOCUS G33851 338 bp DNA linear STS 01-JUL-1997  
DEFINITION human STS SHGC-50253, sequence tagged site.  
ACCESSION G33851  
VERSION G33851.1 GI:2227155  
KEYWORDS STS; STS sequence; primer; sequence tagged site.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 338)  
AUTHORS Myers, R.M.  
JOURNALS Unpublished (1997)  
COMMENT  
Contact: Richard M. Myers  
Stanford Human Genome Center (SHGC)  
Stanford University School of Medicine  
Department of Genetics, M-344, Stanford, CA 94305, USA  
Tel: 415/7259687  
Fax: 415/7259689  
Email: myers@shgc.stanford.edu  
Primer A: TTACGGCAATAATGGAACCTGC  
Primer B: TGGTACTGAAATGCGACCA  
STS size: 99  
PCR Profile:  
Initial incubation: 94 degrees C for 90 seconds  
Denaturation: 94 degrees C for 15 seconds  
Annealing: 62 degrees C for 23 seconds  
Polymerization: 72 degrees C for 30 seconds  
PCR Cycles: 30  
Thermal Cycler: Perkin Elmer 9600  
Protocol:  
Template: 25 ng  
Primer: each 1 uM  
dNTPs: each 200 uM  
Tag Polymerase: 0.05 units/ul  
Total Vol: 10 ul  
Buffer:  
MgCl2: 2.5 mM  
KCl: 50 mM  
Tris-HCl: 20 mM  
pH: 8.3  
Redeveloped chromosome 4 markers for scoring on oligonucleotide arrays. Designed and developed at the Stanford Human Genome Center on sequences previously developed as STS at the Stanford Human Genome Center or the Whitehead Institute.  
FEATURES  
Source  
1..338  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="4"  
STS  
223..321  
primer\_bind  
223..243

primer\_bind complement(302..321)  
BASE COUNT 101 a 57 c 63 g 114 t 3 others  
ORIGIN  
Query Match 12.6%; Score 329.2; DB 11; Length 338;  
Best Local Similarity 98.2%; Pred. No. 2e-68;  
Matches 331; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
Qy 1129 ATTTCCTATTTTTCATTTTCATATGCTTTTCTGCTAGATTTCAGAGATTTTAA 1188  
Db 2 AATTCCTATTTTTCATTTTCATATGCTTTTCTGCTAGATTTCAGAGATTTTAA 61  
Qy 1189 AATCAGAAATCATCTAAATGATATTAGCCTATCAAAACTCCAAAGCCACACAC 1248  
Db 62 AATCAGAAATCATCTAAATGATATTAGCCTATCAAAACTCCAAAGCCACACAC 121  
Qy 1249 CAGTCCGACATCAATTTGAAAGGTTTTCAGAGAGGAGATTAAGTAGTGAGT 1308  
Db 122 CAGTCCGACATCAATTTGAAAGGTTTTCAGAGAGGAGATTAAGTAGTGAGT 181  
Qy 1309 TTAATTAAGCTTAACCTAGGTAAGTAATGAAATATTAGGCAATTAAGTAAGT 1368  
Db 182 TTAATTAAGCTTAACCTAGGTAAGTAATGAAATATTAGGCAATTAAGTAAGT 241  
Qy 1369 GCTTCAGCTGTTCTTGTTGTTGACTTCTCACTCTAATGTTTAAAGAGCAACAAAGCTTA 1428  
Db 242 GCTTCAGCTGTTCTTGTTGTTGACTTCTCACTCTAATGTTTAAAGAGCAACAAAGCTTG 301  
Qy 1429 TGGTCCATTTTCAGTAACACGCGTGTGTTTAGATG 1465  
Db 302 TGGTCCATTTTCAGTAACACGCGTGTGTTTAGATG 338  
RESULT 10  
LOCUS AX198087 470 bp DNA linear PAT 30-AUG-2001  
DEFINITION Sequence 542 from Patent WO0151513.  
ACCESSION AX198087  
VERSION AX198087.1 GI:15388402  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 470)  
AUTHORS Algate, P.A.  
TITLE Ovarian tumor-associated sequences  
JOURNAL Patent: WO 0151513-A 542 19-JUL-2001;  
CORIXA CORPORATION (US)  
FEATURES  
Source  
1..470  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 119 a 98 c 96 g 151 t 6 others  
ORIGIN  
Query Match 11.6%; Score 303.2; DB 6; Length 470;  
Best Local Similarity 97.9%; Pred. No. 3.7e-62;  
Matches 327; Conservative 0; Mismatches 5; Indels 2; Gaps 2;  
Qy 1952 CCTGATTTTGTGCTTATTCACACTAAGTAGATCATCTCTTCTTTTAAAGTTAA 2011  
Db 2 CCTGATTTTGTGCTTATTCACACTAAGTAGATCATCTCTTCTTTTAAAGTTAA 61  
Qy 2012 TGAGGAATTTAGCTCCACATCAACAAATGCAATTCAGCAAGGCTTCATATCTCTGCT 2071  
Db 62 TGAGGAATTTAGCTCCACATCAACAAATGCAATTCAGCAAGGCTTCATATCTCTGCT 121  
Qy 2072 GTGGCTCGTGAATAAGAGAGCTTATTCAGATTTCCTGCCCTAGCTTTAAGCTCCACTTAC 2131  
Db 122 GTGGCTCGTGAATAAGAGAGCTTATTCAGATTTCCTGCCCTAGCTTTAAGCTCCACTTAC 181  
Qy 2132 ATGCTGGAAGACTGGCGTGAAGGACAGATGATTCATCTGCTGTTACTGAAAAACAGGTGT 2191

```

Db 182 ATGCTGAGACCGGCGTAGGACAGATGTAATCTGCGTAGTGAAGAAAGAGGTGT 241
Qy 2192 GATCCTGTAGTACTATAAGAGACCTAAATGTCATGTTCAATTTAGCAAGTCTTC 2251
Db 242 GATCCTGTAGTACTATAAGAGACCTAAATG -CAGTGTCAATTTAGCCAGNG-TC 299
Qy 2252 TAACAACTAACTCTTCAATGCTTGAAGAGAT 2285
Db 300 TAACAACTAACTCTTCAATGCTTGAAGAGAT 333

RESULT 11
AX208670 470 bp DNA linear PAT 31-AUG-2001
DEFINITION Sequence 510 from Patent WO0157207.
AX208670
VERSION AX208670.1 GI:15423093
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE Algate, P.A. and Mannion, J.
JOURNAL Compositions and methods for the therapy and diagnosis of ovarian
PATENT: WO 0157207-A 510 09-AUG-2001;
CORIXA CORPORATION (US)
FEATURES
source location/Qualifiers
1..470
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 119 a 98 c 96 g 151 t 6 others
ORIGIN
Query Match 11.6%; Score 303.2; DB 6; Length 470;
Best Local Similarity 97.9%; Pred. No. 3.7e-62;
Matches 327; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

Qy 1952 CCGATTTTGTCTTATTCACACTAAGTAGATCATTCCTTTTATGTTAA 2011
Db 2 CCGATTTTGTCTTATTCACACTAAGTAGATCATTCCTTTTATGTTAA 61
Qy 2012 TGAGAGATTTAGCCTCCACTCAACATGTTCAATTCAGCAAGCTTTCATCTTCT 2071
Db 62 TGAGAGATTTAGCCTCCACTCAACATGTTCAATTCAGCAAGCTTTCATCTTCT 121
Qy 2072 GTGGGCTGCTGATTAAGAGCTTATTCAGGTTTCTGCTAGCTATTAGCTCCACTTAC 2131
Db 122 GTGGGCTGCTGATTAAGAGCTTATTCAGGTTTCTGCTAGCTATTAGCTCCACTTAC 181
Qy 2132 ATGCTGAGAGCTGGCGTAGGAGACAGATGTAATCTGCTGTTAGTGAAGAGGTGT 2191
Db 182 ATGCTGAGAGCTGGCGTAGGAGACAGATGTAATCTGCTGTTAGTGAAGAGGTGT 241
Qy 2192 GATCCTGTAGTACTATAAGAGACCTAAATGTCATGTTCAATTTAGCAAGTCTTC 2251
Db 242 GATCCTGTAGTACTATAAGAGACCTAAATG -CAGTGTCAATTTAGCCAGNG-TC 299
Qy 2252 TAACAACTAACTCTTCAATGCTTGAAGAGAT 2285
Db 300 TAACAACTAACTCTTCAATGCTTGAAGAGAT 333

RESULT 12
G21512/C 425 bp DNA linear STS 31-MAY-1996
LOCUS G21512
DEFINITION human STS WI-14698, sequence tagged site.
ACCESSION G21512
VERSION G21512.1 GI:1341838
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE human STS derived from sequences in dbEST and the Unigene

```

```

ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE Algate, P.A. and Mannion, J.
JOURNAL Compositions and methods for the therapy and diagnosis of ovarian
PATENT: WO 0157207-A 510 09-AUG-2001;
CORIXA CORPORATION (US)
FEATURES
source location/Qualifiers
1..425
/organism="Homo sapiens"
/db_xref="taxon:9606"
/mab="532.1 CR from top of Chr4 linkage group"
STS
primer_bind 16..40
primer_bind complement(120..144)
BASE COUNT 135 a 84 c 87 g 112 t 7 others
ORIGIN
Query Match 8.5%; Score 221; DB 11; Length 425;
Best Local Similarity 90.7%; Pred. No. 2.1e-42;
Matches 302; Conservative 0; Mismatches 22; Indels 9; Gaps 6;

Qy 2249 TTCTAACAACTAACTCTTCAATGCTTGAAGAGATAC -TACAAAGCAATCTTTA 2304
Db 333 TCCTAACAACTAACTCTTCAATGCTTGAAGAGATAC -TACAAAGCAATCTTTA 274
Qy 2305 TAGAATTTGGGCGC -TAGAATTTGGGCGC -GCTATGTTAAGCTCCAAA 2361
Db 273 TAGAATTTGGGCGC -TAGAATTTGGGCGC -GCTATGTTAAGCTCCAAA 214
Qy 2362 GCTTCACTGTGTTT -CTGCCGCTG -TCCTGAGTTGTCACCACTGACTGGGCAAGCTTC 2419
Db 213 GCTTCACTGTGTTT -CTGCCGCTG -TCCTGAGTTGTCACCACTGACTGGGCAAGCTTC 154
Qy 2420 TTGGGCACTGATGTAAGAGCTTGTCTTTTCCACTTAACAGTTATCTTTGACTCTTC 2479
Db 153 TTGGGCACTGATGTAAGAGCTTGTCTTTTCCACTTAACAGTTATCTTTGACTCTTC 94

```

QY 2480 CCTGTTATGCTTACAAAATGCTGATGCTTATGGAAGGCTGTTAAATTAATTCCTGTT 2539  
 |||  
 Db 93 CCTGTTATGCTTACAAAATGCTGATGCTTATGGAAGGCTGTTAAATTAATTCCTGTT 34  
 QY 2540 AAAGAAATTAAGTTGCTATTTTGGACAT 2572  
 |||  
 Db 33 AAAGAAATTAAGTTGCTATTTTGGACAT 1

RESULT 13  
 G33614/c 425 bp DNA linear STS 01-JUL-1997  
 LOCUS G33614  
 DEFINITION human STS SHGC-50534, sequence tagged site.  
 ACCESSION G33614  
 VERSION G33614.1 GI:2226918  
 KEYWORDS STS: STS sequence; primer: sequence tagged site.  
 SOURCE human  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 425)  
 AUTHORS Myers, R.M.  
 JOURNAL Unpublished (1997)  
 COMMENT

Contact: Richard M. Myers  
 Stanford Human Genome Center (SHGC)  
 Stanford University School of Medicine  
 Department of Genetics, M-344, Stanford, CA 94305, USA  
 Tel: 4157259687  
 Fax: 4157259689  
 Email: myers@shgc.stanford.edu

Primer A: TAATTTACAGCCTTCATTAAGCC  
 Primer B: ACNAACTAACTCCTCAATGCC  
 STS size: 284  
 PCR Profile:  
 Initial incubation: 94 degrees C for 90 seconds  
 Denaturation: 94 degrees C for 15 seconds  
 Annealing: 62 degrees C for 23 seconds  
 Polymerization: 72 degrees C for 30 seconds  
 PCR Cycles: 30  
 Thermal Cycler: Perkin Elmer 9600  
 Protocol:  
 Template: 25 ng  
 Primer: each 1 uM  
 dNTPs: each 200 uM  
 Tag Polymerase: 0.05 units/uL  
 Total Vol: 10 uL

Buffer: MgCl2: 2.5 mM  
 KCl: 50 mM  
 Tris-HCl: 20 mM  
 pH: 8.3

Redeveloped chromosome 4 markers for scoring on oligonucleotide arrays. Designed and developed at the Stanford Human Genome Center on sequences previously developed as STSs at the Stanford Human Genome Center or the Whitehead Institute.

FEATURES  
 Location/Qualifiers  
 source 1..425  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /map="4"  
 STS 45..328  
 primer\_bind 45..68  
 primer\_bind complement(305..328)  
 BASE COUNT 135 a 84 c 87 g 112 t 7 others  
 ORIGIN

Query Match 8.5%; Score 221; DB 11; Length 425;  
 Best Local Similarity 90.7%; Pred. No. 2.1e-42;

Matches 302; Conservative 0; Mismatches 22; Indels 9; Gaps 6;  
 QY 2249 TTCTACAAACCTAACTC--TTCAATGCTTGGAAAGATAC--TACAAAGCAATCTTTA 2304  
 |||  
 Db 333 TCCTAACCAACTAACTCCTTCAATGCTTGGAAAGATACCTACAAAGCCAACTTTA 274  
 QY 2305 TAGAATTTGGGCC--AAGATAAATCTATGTTGTTTGCATG--GCTATTTGAAGTCCAAA 2361  
 |||  
 Db 273 TAGAATTTGGGCCCAAGATAAATGATCAATGTTTGGCATGTCATTTGTAAGTCCAAA 214  
 QY 2362 GGTTCACGTGTTT--CTGCGCGTG--TCCTGAGTTGTACACCTACCTGAGGCAAGGCTTC 2419  
 |||  
 Db 213 GGTTCACGTGTTTNCCTGCGCGTGNNCTGAGTTGTACACCTACCTGAGGCAAGGCTTC 154  
 QY 2420 TTGGGCATGATGTGAACTGTGTCTCTTTTCCCACTAACTATCTTTGACTCTTTG 2479  
 |||  
 Db 153 TTGGGCATGATGTGAACTGTGTCTCTTTTCCCACTAACTATCTTTGACTCTCTTG 94  
 QY 2480 CCTGTTATGCTTACAAAATGCTGATGCTTATGGAAGGCTGTTAAATTAATTCCTGTT 2539  
 |||  
 Db 93 CCTGTTATGCTTACAAAATGCTGATGCTTATGGAAGGCTGTTAAATTAATTCCTGTT 34  
 QY 2540 AAAGAAATTAAGTTGCTATTTTGGACAT 2572  
 |||  
 Db 33 AAAGAAATTAAGTTGCTATTTTGGACAT 1

RESULT 14  
 AX015378 1491 bp DNA linear PAT 07-SEP-2000  
 LOCUS AX015378  
 DEFINITION Sequence 40 from Patent WO9951727.  
 ACCESSION AX015378  
 VERSION AX015378.1 GI:10041358  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1491)  
 AUTHORS Schmitt, A., Specht, T., Dahl, E., Hinzmann, B., Rosenthal, A. and  
 Pilarsky, C.  
 TITLE Human nucleic acid sequences of normal ovary tissue  
 JOURNAL Patent: WO 9951727-A 40 14-OCT-1999;  
 SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN  
 BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN  
 (DE); PILARSKY CHRISTIAN (DE)  
 FEATURES  
 Location/Qualifiers  
 source 1..1491  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 323 a 362 c 455 g 351 t  
 ORIGIN

Query Match 2.6%; Score 67.2; DB 6; Length 1491;  
 Best Local Similarity 51.2%; Pred. No. 2.3e-05;  
 Matches 206; Conservative 0; Mismatches 193; Indels 3; Gaps 2;  
 QY 351 GAGGAATCAACTGGCCAGATTTGGTTAAATGCTGGAAGAACTGTCTGTCCAATCAAG 410  
 |||  
 Db 194 GATGAACACTTGTGTGCCAAGCTGATGACGTCTGCAAGAGACAGCTGGCCAGGCGGG 253  
 QY 411 CAAACTAAACTTGTGTCTCAAAAGTCTTGTCTGAGAAACTGACCCAGAGAAATGCT 470  
 |||  
 Db 254 CTGGGCTCTCGACGCCCTGCGGCTGCTGATGCTTAGCCACTGTGTAAGCCAGGTGGG 313  
 QY 471 CAAAGATGTCCTGCGCTTTCCTCAACGAGAGCCCTGCGGCTTTCGAGGTGTGTATGAC 530  
 |||  
 Db 314 AAAGACTACTAGCTGCGCTGACGAGCGGCTGCGGCTGCGGCGGCTGCTGAGC 373  
 QY 531 GTGACTTGGAATTAAGAAATGTATGTAAGAAAGCTGATAGATGTGTGTATTTCTAGC 590  
 |||  
 Db 374 GTCTCGTGTGACAGGCAAGAGCTGACAGCGTGGGCGCACTGGCACTGACCCAGC 433

QY 591 GTGTAACCTACTTTTGTAGCTTGTGTTAAGCAGA-GAAGTGTCTAGACTAG 649  
 Db 434 CTGTGCGCCACCTTCCAGCTGACCTGCTGCGCTGACACACACTGCGCCCAAG 493  
 QY 650 CTTCAGGACTTTT--CTTACTAGAGCTGCTTCTCTGTTTCAAGAGAAGCTGTG 707  
 Db 494 ATCCAGGGGCTGTTTGTGCTCCGCAACTCTCCCTCTGCTGCTGCTGCTGCTGCTG 553  
 QY 708 ATCCAGCTCAGACTTGTGCTGTTAAGAAAACTTAC 749  
 Db 554 AGCGTAGCAGCTGCTTCCGAGTCATCAGAGAGAGCTGTAC 595

## RESULT 15

BC015236

DEFINITION BC015236 1732 bp mRNA linear PRI 29-OCT-2001  
 Home sapiens, hypothetical protein, clone MGC:17872 IMAGE:3906374,  
 mRNA, complete cds.

ACCESSION BC015236

VERSION BC015236.1 GI:15929605

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REMARK

## COMMENT

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
 Contact: MGC help desk  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www.shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcddpaxil.stanford.edu](mailto:mcddpaxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
 R. M.

## FEATURES

## source

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>  
 Series: IRAC Plate: 14 Row: p Column: 1.  
 Location/Qualifiers  
 1. 1732

## CDS

/organism="Homo sapiens"  
 /db\_xref="locusID:54541"  
 /db\_xref="taxon:9606"  
 /clone="MGC:17872 IMAGE:3906374"  
 /tissue\_type="uterus, leiomyosarcoma"  
 /clone\_id="NIH\_MGC\_71"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV-SPORT6"  
 177..875  
 /codon\_start=1  
 /product="hypothetical protein"  
 /protein\_id="AAH15236.1"  
 /db\_xref="GI:15929606"  
 /translation="MPSIMDRFSSSTSSSPSLPTPTFDPRPSAWGSATREGEFD  
 RSTSLSSDCESLDSNSGPEEDTAYLDVGSIPDELSDPEHLICANIMLLOE  
 SLAORIGSRPARLLPMSOLVOYGEKELRLAYSEKGLRGALIDVCYDQGSCHSV  
 GOLADPSLVPRTQTLVLRDLDRMPKIDGLFSSANSFPLPGFSOSLTSTGFRVIV  
 KRIYSSQQLIEEC"

## BASE COUNT

347 a 459 c 536 g 390 t

Query Match 2.68; Score 67.2; DB 9; Length 1732;  
 Best Local Similarity 51.28; Pred. No. 2.4e-05;

Matches 206; Conservative 0; Mismatches 193; Indels 3; Gaps 2;  
 QY 351 GAGGAATCAACTTGCAGAAATTTGGTTAAATGCTGAGAACTGTCTGCTCAATCAAG 410  
 Db 441 GATGAACACTTGTGTGCAACCTGATGCTGCTGAGAGAGCTGCGCCAGCCGG 500  
 QY 411 CAACACTAACTTGTGCTTAAAGTCTTGTCTGCTGAGAAACTGACGAGAAATGCT 470  
 Db 501 CTGGGCTCTGAGCCCTGCGCGCTGATGCTGAGCAAGTGTGAAGCCAGGTGGC 560  
 QY 471 CAAGATGCTCTGCGGCTTCTCAACGAGCCCGGCGCTGAGAGTGTGATATCAC 530  
 Db 561 AAGAACTACTGGCGCTGCGCTACAGGAGCCGTGCGGGGGCGCTGCTGAC 620  
 QY 531 GTCAACTTGAATAATGAATAATGATGAATAAGCTGATGATGATGATGATGATGATG 590  
 Db 621 GTCTGCTGAGAGGAGGCAAGAGCTGCGCACAGGTGGCCAGCTGGCAGTGGCAGC 680  
 QY 591 GTGTAACCTACTTTTGTAGCTTGTGTTAAGCAGA-GAAGTGTCTAGACTAG 649  
 Db 681 CTGTGCGCCACCTTCCAGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 740  
 QY 650 CTTCAGGACTTTT--CTTACTAGAGCTGCTTCTCTGTTTCAAGAGAAGCTGTG 707  
 Db 741 ATCCAGGGGCTGTTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 800  
 QY 708 ATCCAGCTCAGACTTGTGCTGTTAAGAAAACTTAC 749  
 Db 801 AGCGTAGCAGCTGCTTCCGAGTCATCAGAGAGAGCTGTAC 842

Search completed: October 13, 2002, 03:03:37.  
 Job time : 4759 secs

